

6. FOLGE
UNIVERSITY
OF MICHIGAN

1950, BAND 8, HEFT 1-2

NOV 27 1950

PHYSICS
LIBRARY

ANNALEN DER PHYSIK

Gegründet 1799 durch F. A. C. Gren
Fortgeführt durch L. W. Gilbert, J. C. Poggendorff,
G. und E. Wiedemann, P. Drude,
W. Wien, M. Planck, E. Grüneisen

Der ganzen Reihe 443. Band

Kuratorium:

W. GERLACH, CHR. GERTHSEN, F. HUND, W. KOSSEL,
M. v. LAUE, W. MEISSNER, R. W. POHL, R. ROMPE,
R. SEELIGER, A. SOMMERFELD, W. WEIZEL

Herausgegeben von

F. MÖGLICH
Berlin



JOHANN AMBROSIOUS BARTH / VERLAG / LEIPZIG

Redaktionssechluß am 31. Juli 1950

Inhalt

	Seite
Hans Wolter, Kiel: Zweidimensionale Farbschlierenverfahren. (Mit 6 Abbildungen und einer mehrfarbigen Tafel)	1
Hans Wolter, Kiel: Physikalische Begründung eines Farbenkreises und Ansätze zu einer physikalischen Farbenlehre. (Mit 3 Abbildungen).....	11
Irmgard Wolf, Potsdam: Untersuchungen über die Beeinflussung der spektralen Empfindlichkeit von Selen-Photoelementen durch Photostrom und Außenwiderstand. (Mit 7 Abbildungen)	30
A. Sommerfeld und F. Bopp, München: Zum Problem der Maxwellschen Spannungen	41
A. Sommerfeld und E. Ramberg, München: Das Drehmoment eines permanenten Magneten im Felde eines permeablen Mediums	46
Johannes Fischer, Karlsruhe: Zur Definition der magnetischen Größen	55
M. Czerny, W. Kofink und W. Lippert, Frankfurt a. M., Bolometer geringer Trägheit. (Mit 6 Abbildungen)	65
Wolfgang Böer, Berlin: Über die Beschneidung des Rauschspektrums bei Schwingungskreisen. (Eine Ergänzung zu einer Arbeit von R. Feldtkeller.) (Mit 4 Abbildungen)	87

Für diese Zeitschrift werden grundsätzlich nur Arbeiten angenommen, die vorher weder im Inland noch im Ausland veröffentlicht worden sind. Mit der Annahme des Manuskriptes und seiner Veröffentlichung geht das alleinige Recht der Vervielfältigung, Verbreitung und Übersetzung einschließlich des Rechts der Veranstaltung von Sonderdrucken auf den Verlag über. Von jeder Arbeit liefert der Verlag 75 Sonderabzüge kostenlos.

Es wird gebeten, Manuskripte an Herrn Prof. Dr. F. Möglich, (1) Berlin-Buch, Lindenberger Weg 74, einzusenden.

Die Redaktion bittet, die Manuskripte *druckfertig* einzuliefern und in den Korrekturen den bisherigen Raum des Textes nicht zu überschreiten. Die *Zeichnungen* sind in sorgfältiger Ausführung auf besonderen Blättern beizulegen.

Bei den *Zitaten* wolle man die Abkürzungen der Zeitschriftentitel nach dem Verzeichnis wählen, das jedem Jahrgang der „Physikalischen Berichte“ vorangestellt ist.

Die Zeitschrift erscheint in Bänden mit 8 Heften, die zwanglos ausgegeben werden. Die Lieferung erfolgt durch den Sortimentsbuchhandel, in Ausnahmefällen unmittelbar vom Verlag. Der Bezugspreis beträgt je Band DM. 34.— zuzügl. Postgebühren. Einzelpreis dieses Heftes DM 10.20.

Die Lieferung erfolgt bis zur Abbestellung, die nur für das Ende eines Bandes ausgesprochen werden kann.

Zweidimensionale Farbschlierenverfahren¹⁾

Von Hans Wolter

(Mit 6 Abbildungen und einer mehrfarbigen Tafel)

Inhaltsübersicht

Die Ablenkungen des Lichtes durch ein Schlierenobjekt bilden eine zweidimensionale Mannigfaltigkeit; sie lassen sich daher nicht durch eine eindimensionale Mannigfaltigkeit, wie es die Helligkeit ist, eindeutig und stetig kennzeichnen. Daher kann kein nur mit Hell und Dunkel und der kontinuierlichen Grauskala arbeitendes Schlierenverfahren in einem einzigen Bilde einen ausreichenden Überblick über alle von einem Schlierenobjekt bewirkten Ablenkungen des Lichtes geben. Doch läßt sich dieses Ziel erreichen, wenn außer der Helligkeit noch die Farbe als Kennzeichnungsmittel herangezogen wird. Ein besonders übersichtliches Verfahren — unter mehreren hier diskutierten — entsteht aus dem bekannten abbildenden Schlierenverfahren nach Toepler, wenn die Schlierenblende durch eine in farbige Sektoren nach Art eines Farbenkreises aufgeteilte Filterplatte ersetzt wird. Nicht ablenkende Objektteile werden dann farblos abgebildet, ablenkende erscheinen in einer um so gesättigteren Farbe, je stärker sie das Licht ablenken; die Ablenkrichtung äußert sich in dem Farbton. Das Verfahren bewährte sich bei der Untersuchung von durchsichtigen Objekten auf Schlieren jeder Art, vor allem aber bei der mikroskopischen Beobachtung des Kristallwachstums und in seiner Anwendung auf das Metallmikroskop.

1. Aufgabe und Lösungsprinzip

Im Metallmikroskop kann man durch einen Eingriff in den Strahlengang bekanntlich ähnlich wie bei dem Toeplerschen Schlierenverfahren die verschiedenen Flächen der Kristalliten verschieden hell abbilden; die jeweilige Helligkeit, mit der eine Fläche im Bilde erscheint, läßt dann einen Schluß auf die Richtung der Flächennormalen zu.

Man kann jedoch ganz allgemein einsehen, daß die Flächennormalenrichtung durch die — eindimensionale — Helligkeitsskala nicht eindeutig gekennzeichnet werden kann, wie geschickt das Verfahren im einzelnen auch durchgeführt sein möge; denn die Spitzen aller Flächennormalen (Einheitsvektoren) zu den durch einen festen Raumpunkt gehenden Ebenen erfüllen eine Kugeloberfläche oder doch wenigstens einen Teil derselben (wenn wir den „Meßbereich“ einschränken, wie die praktisch vorkommenden Objekte das meist tun), auf jeden Fall also eine zweidimensionale Mannigfaltigkeit.

Nun kann man zwar prinzipiell ein zweidimensionales Kontinuum auf ein eindimensionales eineindeutig „abbilden“ (dieses Wort ist hier im Sinne der Mengenlehre gebraucht — man nennt das spezielle Verfahren ein „Diagonalverfahren“); jedoch ist die Abbildung grundsätzlich für jeden Punkt des Kontinuums unstetig und daher für naturwissenschaftliche Zwecke unbrauchbar, da noch so kleine Beobachtungsfehler stets große Fehler an der zu messenden Größe bewirken würden.

¹⁾ Vorgetragen auf der Tagung der Deutschen Physikalischen Gesellschaft in Bonn am 29. September 1949.

Da es also viel mehr zu kennzeichnende Flächennormalen gibt, als kennzeichnende Hilfsmittel zunächst verfügbar sind, wird ein weiteres Mittel der Kennzeichnung erforderlich. Als solches sei hier die Farbe benutzt. Diese ist schon früher zu Farbschlierenverfahren herangezogen worden²⁾, aber nicht zu dem hier in Rede stehenden Zweck, sondern nur zu einer gegenüber dem Toeplerschen Verfahren augenfälligeren eindimensionalen Kennzeichnung. Da die Farben — z. B. die des Spektrums — ein von der Helligkeit unabhängiges Kontinuum bilden, ist prinzipiell durch Kombination von Helligkeit und Farbe ein zur Kennzeichnung ausreichendes Kontinuum zu erhalten, das im Unterschied zu der reinen Helligkeit ein zweidimensionales sein kann.

Die am Metallmikroskop als Beispiel diskutierte Aufgabe wollen wir auf die Abbildung von Schlierenobjekten allgemein übertragen. Liegt z. B. als Objekt eine Platte homogenen Glases vor, deren beide Grenzflächen uneben sind, so wird freilich kein abbildendes Schlierenverfahren in einem Bilde über die Abweichungen beider Grenzflächen von Ebenen Auskunft geben können. Aufgabe im Sinne dieser Arbeit sei daher im allgemeinen nur die Feststellung der von verschiedenen Gebieten eines Schlierenobjekts bewirkten Lichtablenkungen nach beiden Dimensionen, z. B. Betrag und Richtung, zugleich, mit Hilfe einer einzigen Aufnahme. Für veränderliche Vorgänge wie das Wachsen von Kristallen ist das besonders wichtig, da zwei Aufnahmen desselben Zustandes nacheinander — zur Feststellung etwa zweier Komponenten der Lichtablenkung — meist nicht möglich sind. Das oben angegebene Lösungsprinzip bleibt auch für diese verallgemeinerte Aufgabe unverändert.

2. Spezielle Geräte zur Realisierung des Lösungsprinzips

a) Prismenverfahren

Die oben genannte grundsätzliche Lösung des Problems gewinnt apparative Gestalt am einfachsten durch Abwandlung des von H. Schardin³⁾ beschriebenen Prismenfarbschlierenverfahrens, indem man statt des dort benutzten Spaltes ein annähernd punktförmiges Lichtquellendiaphragma verwendet und in die Ebene, die dort den Steg als Schlierenblende enthält, zusätzlich einen Graukeil mit einem zum Steg parallelen Opazitätsgradienten stellt. Solche leicht herstellbaren Apparaturen werden jedoch übertriften von der in Abb. 1 skizzierten, da diese außerdem von den speziellen zyklischen Eigenschaften des Farbenkreises Gebrauch macht; die Wirkungsweise sei daher speziell unter Bezug auf diese eine Apparatur beschrieben und für andere später ergänzt, soweit das noch erforderlich scheinen wird.

b) Polarkoordinatenverfahren im Hellfeld

In Abb. 1 bezeichnet *a* eine von links beleuchtete Kreisblende, die aus der optischen Achse gesehen einen Anblick nach Abb. 2a zeigen würde. Sie befindet sich in der lichtseitigen Brennebene des Kondensors *b*, der jedes von einem Punkte der Blende *a* ausgehende Lichtbündel in ein Parallelstrahlenbündel umwandelt. Auf dem Objektisch *c* liegt ein durchsichtiges Schlierenobjekt, z. B. ein Objektträger mit aus der Mutterlauge auskristallisierenden Kristallen oder eine Glasplatte mit mehr oder weniger großen Abweichungen von der Planparallelität der Flächen oder von der Homogenität des Materials.

Das Objektiv *d* bildet das Objekt auf die Beobachtungsebene *f* ab, und das so entstehende Bild gibt ohne weiteres keinen Aufschluß über die Schlieren, Kristall-

²⁾ H. Schardin, Ergebnisse der exakten Naturwissenschaften, XX, 303 (1942).

³⁾ ebenda S. 346.

flächen o. ä. Doch genügt bekanntlich die Anbringung einer Messerschneide in der bildseitigen Brennebene *e* des Objektivs nach dem Vorgange Toeplers, um

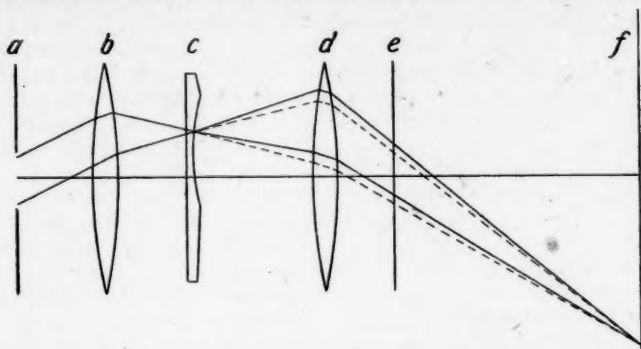


Abb. 1. Zur Abbildung eines Schlierenobjekts *c* mit zweidimensionalem Farbschlierenverfahren; *a* Blende, *b* Kondensor, *d* Objektiv, *e* Farbsektorplatte, *f* Beobachtungsebene

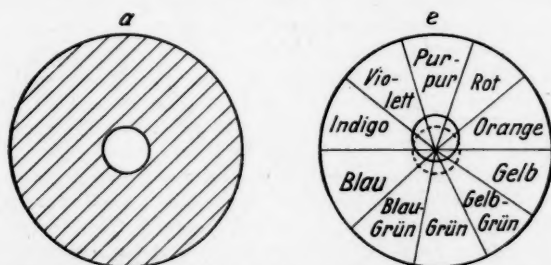


Abb. 2. Aufsichtsbilder der Blende *a* und der Farbsektorplatte *e*, letzte mit unabgelenktem (gestrichelt) und abgelenktem (ausgezogen) Bilde der Blende

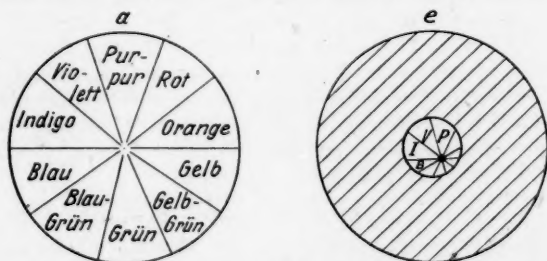


Abb. 3. Zum zweidimensionalen Farbschlierenverfahren nach Vertauschen von Blende und Farbsektorplatte

die Schlieren durch Hell- und Dunkelschattierung im Bilde sichtbar zu machen. Die oben genannten prinzipiellen Mängel dieses (und eines jeden nur mit Helligkeitsunterschieden arbeitenden) Schlierenverfahrens werden nun dadurch behoben,

daß in der bildseitigen Brennebene des Objektivs e statt der Toeplerschen Schneide eine Farbsektorplatte angebracht wird, wie Abb. 2e sie aus Richtung der optischen Achse zeigt. Die Farben seien z. B. durch Anfärben einer Gelatinefolie (einer ausfixierten unentwickelten Photoplatte) mit Anilinfarben hergestellt und mögen etwa dem Ostwaldschen Farbenkreise ⁴⁾ entsprechen.

Solange kein Objekt auf dem Objektisch c liegt, wird die Kondensorblende a durch den Kondensor b und das Objektiv d scharf in die Ebene e als Kreis abgebildet. Die Farbsektorplatte sei symmetrisch zu diesem Kreise justiert, so daß durch die Farbmischung das gesamte Gesichtsfeld der Beobachtungsebene f gleichmäßig weiß erscheint.

Legt man nun auf den Objektisch c z. B. eine Glasplatte, deren begrenzende Ebenen einen Keilwinkel einschließen, und lenkt diese Platte z. B. das Licht infolge dieses Keilwinkels nach oben ab, so wird das Kreisbild der Kondensorblende nun auf der Farbsektorplatte nach oben verschoben, und das Licht passiert daher einen im Mittel purpurfarbigen Bereich, wie die Abb. 1 und 2e das andeuten. Die prismatische Glasplatte wird also in diesem Falle purpurfarben auf der Beobachtungsebene f abgebildet werden, und zwar wird die Farbe um so gesättigter erscheinen, je mehr das Blendenbild aus der Symmetrielage verschoben wurde, d. h. je stärker die Glasplatte das Licht aus seiner ursprünglichen Richtung (in Abb. 1 gestrichelt) ablenkt.

Eine zweite, z. B. nach unten ablenkende Glasplatte, die neben die erste Platte auf den Objektisch gelegt wird, muß in entsprechender Weise grün abgebildet werden, und eine Platte mit Bezirken, die das Licht verschieden ablenken, wird ein vielfarbiges Bild, ergeben, dessen Farbton an jeder Stelle die Richtung und dessen Farbsättigung die Größe jener Ablenkung für die entsprechende Objektstelle kennzeichnet.

Die Aufnahmen Tafel Ia-e veranschaulichen die Wirkung dieses Verfahrens. Die Apparatur benutzte als Kondensor eine Zeiß-Protarlinse VII mit 59 cm Brennweite und als Objektiv ein Voigtländer-Heliostigmat 1:2,5 mit 33,5 cm Brennweite ⁵⁾. Die Tafel Ia-c zeigt einen Glasrichter in Achsenrichtung, schräg bzw. senkrecht zur Achse gesehen; Tafel Id gibt das Bild einer schlechten Glasplatte und Tafel Ie das zweier — teilweise übereinandergelegter — Glasplatten, die dem Verfasser als brauchbare Planparallelplatten angeboten worden waren. Daß diese einen keilförmigen Randbezirk haben, beweist die bläuliche Färbung in unserem Bilde. Wo beide Platten sich überdecken, ist die Farbe vertieft bei unverändertem Farbton; das entspricht dem vergrößerten Ablenkbetrag bei unveränderter Ablenkrichtung im Vergleich mit der Einzelplatte. Da übrigens beide Platten jede für sich betrachtet völlig gleichwertige Bilder ergeben, konnte leicht nachgewiesen werden, daß sie aus einer Platte durch Auseinanderschneiden hergestellt worden waren.

c) Polarkoordinatenverfahren im Dunkelfeld

Da das beschriebene Schlierenverfahren die Ablenkrichtung durch den Farbton und den Ablenkbetrag durch die Farbsättigung kennzeichnet, bewirkt es zwar die geforderte zweidimensionale Erfassung der Ablenkungen in einem Bilde; aber

⁴⁾ Über einen zweckmäßigeren Farbenkreis wird Verfasser auf der Tagung der Deutschen Physikalischen Gesellschaft in Münster berichten.

⁵⁾ Als Aufnahmekamera bewährte sich wegen der ständigen Beobachtbarkeit des Bildes und der Vielseitigkeit des Verschlusses vor allem die Kine-Exakta.

das geschieht in etwas anderer Weise, als es dem im Abschnitt 1 dargelegten Lösungsprinzip entsprechen würde. Wollen wir, wie dort geplant, die Farbe und die Helligkeit statt der Farbe und der Farbsättigung als zweidimensionale Kennzeichnungshilfsmittel benutzen, so kann dazu leicht eine zusätzliche undurchlässige Kreisscheibe in der Mitte der Farbsektorplatte verhelpen. Diese muß so bemessen und justiert sein, daß sie genau das — durch kein Objekt veränderte — Bild des Lichtquellendiaphragmas abdeckt. Im Bilde erscheinen nicht ablenkende Objektteile dann dunkel, ablenkende um so heller, je größer der Ablenkbetrag ist. Wieder gibt die Farbe die Ablenkrichtung an.

Dieses Verfahren verhält sich zu dem in 2b beschriebenen wie ein Dunkelfeldverfahren zu einem Hellfeldverfahren. Sein Vorteil liegt in der auffälligeren Wirkung sehr kleiner Ablenkungen, sein Nachteil nicht nur in den höheren Anforderungen an die Justierung sondern in der Möglichkeit größerer Fehldeutungen bei Vorliegen von absorbierenden oder unsauberen Objekten und vor allem in dem schlechteren Auflösungsvermögen bei der mikroskopischen Anwendung (siehe unter 3!).

3. Anwendung im Mikroskop

Jedes mit Kondensor ausgerüstete Mikroskop ist bereits weitgehend ein Gerät, wie es durch Abb. 1 angedeutet und in Abschnitt 2b beschrieben wurde; ergänzend ist nur noch in der hinteren Brennebene des Objektivs (Ebene der Aperturblende, in der bei dem Phasenkontrastverfahren des Zernikeplättchen liegen würde) eine Farbsektorplatte einzujustieren. Daß in Abb. 1 ein Mikroskopokular nicht berücksichtigt wurde, ist hier ohne Belang.

a) Hellfeldverfahren mit vertauschter Blende und Farbsektorplatte

Die Tafel If-h zeigt mikroskopische Aufnahmen wachsender K_2SO_4 -Kristalle mit einem gegenüber dem unter 2b beschriebenen noch etwas vereinfachten Verfahren. Da eine in die hintere Brennebene des Objektivs zu setzende Farbsektorplatte klein und sehr präzise ausgeführt sein muß, erwies es sich hier als einfacher, die Farbsektorplatte und das Lichtquellendiaphragma ihre Plätze tauschen zu lassen. Der Kondensor hat eine größere Brennweite als das Objektiv, und die Farbsektorplatte in der lichtseitigen Kondensorbrennebene kann größer und dementsprechend gröber sein; vor allem aber braucht man von dem sie an verschiedenen Stellen durchsetzenden Licht keine Kohärenz zu fordern; die Farbsektorplatte kann hier sogar mattiert ausgeführt werden, während sie in der hinteren Brennebene des Objektivs für jede Lichtfrequenz einen über die ganze Platte konstanten Phaseneinfluß haben muß.

Die Wirkungsweise des durch den genannten Platztausch abgewandelten Verfahrens sei an Abb. 3 im Zusammenhang mit Abb. 1 beschrieben. Kondensor und Objektiv (*b* und *d* in Abb. 1) entwerfen von der Farbsektorplatte (*a* in Abb. 3) ein Bild in der hinteren Brennebene des Objektivs; hier befindet sich nun eine Kreisblende, die aus dem ursprünglichen Lichtquellendiaphragma bei dem Platztausch hervorgegangen zu denken ist. Praktisch braucht man bei Objektiven, deren numerische Apertur wesentlich kleiner als die des Kondensors ist, keine Blende mehr in die hintere Brennebene des Objektivs zu setzen, da die stets vorhandene Aperturblende diese Rolle übernimmt.

Im Innern dieser Aperturblende liegt dann ein Teil aus dem Bilde der Farbsektorplatte. Ein nichtablenkender Objektteil erzeugt ein symmetrisch in der Blende liegendes Farbsektorbild und erscheint daher dem Beobachter weiß, da

alle Farben gleichmäßig zum Bilde dieses Objektteiles in der Beobachtungsebene beitragen. Ablenkende Objektteile geben mehr oder weniger unsymmetrisch zur Blende liegende Bilder der Farbsektorplatte (Abb. 3b) und erscheinen auf der Beobachtungsebene daher in einer Farbe, deren Sättigung den Ablenkbetrag und deren Farbton die Ablenkrichtung kennzeichnet.

Blickt man nach Entfernen des Okulars in ein so eingerichtetes Mikroskop, so sieht man innerhalb der Aperturblende das scharfe Bild der Farbsektorplatte, solange sich kein Objekt im Strahlengang befindet. Wird ein Objekt (z. B. ein Objektträger mit wachsenden Kristallen) auf den Objektisch gelegt, so verschwimmt das Bild der Farbsektorplatte um so mehr, je unterschiedlicher die verschiedenen Teile des Objekts das Licht ablenken. Oft gibt auch ein Objektträger mit klarer Flüssigkeit und Deckglas bedeckt bereits eine kleine Verschiebung des Farbsektorbildes (und daher eine schwache einheitliche Gesichtsfeldfärbung bei der Beobachtung mit Okular); dieser Fehler kann durch Nachjustieren der Farbsektorplatte selbst kompensiert werden.

Die Aufnahmen der Tafel If-h wurden mit einem Objektiv $8\times$, num. Apertur $= 0,2$ und einer (der Farbsektorplatte zur Verfügung stehenden) Beleuchtungsapertur $0,96$ des Kondensors durchgeführt und sind durch ein Okular $10\times$ und eine Nachvergrößerung auf die Gesamtvergrößerung $100\times$ gebracht worden. Die leichten Abweichungen der Untergrundfärbung vom reinen Grau liegen an der soeben genannten prismatischen Form des Systems Objektträger + Flüssigkeit + Deckglas; auf ihre Beseitigung wurde verzichtet, da die Halterung der Farbsektorplatte bei dem verfügbaren Mikroskop nicht zentrierbar war.

Wenn das hier benutzte Verfahren auch besonders einfach*) ist, so muß doch auf das oben angegebene (Farbsektorplatte im Objektivtubus) zurückgegriffen werden, sobald starke Vergrößerungen und also große Objektivapertur gewünscht werden. Es liegt im Wesen all dieser Farbschlierenverfahren, daß eine der beiden Aperturen (des Objektivs bzw. des Kondensors) wesentlich kleiner als die andere sein muß. Für ein hohes Auflösungsvermögen ist nun die große Objektivapertur entscheidend und daher nur die ursprünglich diskutierte Anordnung 2b verwendbar.

b) Farbschlierenverfahren als mikroskopisches Dunkelfeldverfahren

Einfacher und besser als mit dem unter 2c beschriebenen Dunkelfeldverfahren kommt man bei der mikroskopischen Anwendung zum Ziel, wenn man einen gewöhnlichen Kardioidkondensor benutzt und in seine untere Brennebene zentriert eine Farbsektorplatte legt; das Mikroskop ist im übrigen genau so zu justieren und zu bedienen wie bei dem gewöhnlichen Dunkelfeldverfahren. Die Wirkung entspricht der unter 2c beschriebenen; doch ist das Auflösungsvermögen besser, da die Objektivapertur voll ausgenutzt wird.

4. Abwandlungsmöglichkeiten des Verfahrens

Ähnlich wie bisher für Polarkoordinaten lassen sich für jedes Koordinatensystem auf der Kugel 4 Farbschlierenverfahren konstruieren. Daraus erhält man meist noch weitere, indem man die durch den Farbton gekennzeichnete Koordinate nun durch die Helligkeit (bzw. Farbsättigung) kennzeichnet und umgekehrt. Doch

*) Vorläufer dieser einfachen Form sind die älteren qualitativen Verfahren, bei denen auch eine mehrfarbige Platte (z. B. mit konzentrischen Farbringen) in die Kondensorblende eingelegt wurde.

scheiden von vornherein viele der so entstehenden Verfahren aus, da die Eindeutigkeit und Stetigkeit nicht gewahrt werden können.

Wollte man z. B. in Polarkoordinaten den Ablenkbetrag durch den Farbton und die Ablenkrichtung durch die Helligkeit kennzeichnen, so wäre gerade das zweite nicht stetig und uneindeutig möglich. Allgemein läßt sich eine zyklische Variable, wie es hier die Ablenkrichtung ist, nicht durch die von Natur aus nicht zyklische Helligkeit oder Farbsättigung uneindeutig und stetig kennzeichnen.

So zeigt sich dann auch bei der näheren Untersuchung, daß mit den in 2b, c und 3a, b beschriebenen Polarkoordinatenverfahren und dem Verfahren nach 2a die interessanten Verfahren erschöpft sind und irgendwelche Abwandlungen nur noch relativ unbedeutenden oder doch keine wesentlichen Vorteile mehr bietenden Verzerrungen der beiden wichtigsten Koordinatensysteme (Cartesische und Polarkoordinaten) gleichkommen.

5. Zur quantitativen Auswertung

Die quantitative Auswertung der Farbschlierenaufnahmen geschieht zweckmäßig durch Vergleich mit der Aufnahme eines Objekts, das einen hinreichend großen Bereich kontinuierlicher und bekannter Ablenkungen hat; zu einer solchen „Normalschlieren“ eignet sich wie auch bei anderen Schlierenverfahren eine Sammel- oder eine Zerstreuungslinse.

Die Lage von Kristallflächen ließ sich in dieser Weise bis auf Winkelfehler von 0,5 bis zu 3° vermessen.

Bei einem solchen Vergleichsverfahren hat man große Freiheit in der Farbwahl für die Farbsektorplatte; diese kann daher auch sehr einfach durch Anfarben einer Gelatineschicht (s. auf S. 4) oder durch Photographie einer farbigen Zeichnung oder dergleichen auf Agfacolorfilm hergestellt werden.

6. Eine verbesserte Farbsektorplatte

Besser und allerdings auch wesentlich umständlicher sind einige physikalische Wege zur Realisierung des Farbkreises mittels Farben dünner Blättchen oder Polarisationsfarben. Als Beispiel beschreiben wir den Aufbau einer Farbsektorplatte, die auf der Rotationsdispersion optisch aktiver Stoffe beruht.

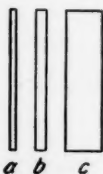


Abb. 4. Farbsektorplatte, auf der Rotationsdispersion beruhend



Abb. 5. Zur Herstellung einer Polarisationsfolie mit ortsabhängiger Polarisationsrichtung

Die Vorrichtung (siehe Abb. 4) besteht aus einem gewöhnlichen Polarisorator *a* zur Herstellung linear polarisierten Lichtes (z. B. einer der bekannten Polaroidfolien), ferner einer Platte *b* aus optisch aktivem Material (z. B. einer Quarzplatte, senkrecht zur Achse geschnitten) und schließlich einem Analysator *c*, den man sich

aus einer Cellopolarfolie folgendermaßen entstanden denken kann. Ein halbkreisförmiges Stück der Cellopolarfolie (Abb. 5d) wird unter Streckung des Materials so zu einem Vollkreise gebogen, daß die ursprünglich parallel zueinander laufenden Polarisationslinien (ausgezogene Linien in Abb. 5d) zu Parabeln (Abb. 5c, Grundriß des Analysators, der in Abb. 5c im Aufriß angedeutet ist) werden. Die Streckung hat dabei so zu geschehen, daß nach ihrer Ausführung jeder vom Mittelpunkt der entstehenden Kreisscheibe gezogene Radiusvektor alle Polarisationslinien unter gleichem Winkel schneidet.

Statt die Polarisationsfolie zu strecken, kann man sie auch ohne Streckung oder Stauchung so zu der Mantelfläche eines Kegels aufrollen, daß die beiden den Halbkreis 6d begrenzenden Durchmesserhälften zu einer Mantellinie des Kegels verschmelzen.

Die gesamte Vorrichtung nach Abb. 4 wirkt, wenn man von rechts (oder von links) her durch sie auf eine mit natürlichem, weißem Licht beleuchtete Fläche sieht, wie eine Farbsektorplatte der Abb. 2e. Jedoch gehen die Farben kontinuierlich ineinander über und sind physikalisch eindeutig reproduzierbar.

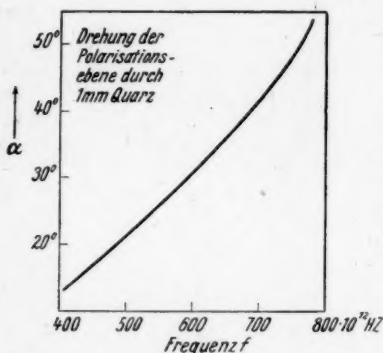


Abb. 6. Rotationsdispersion von Quarz

Die Wirkungsweise ist folgendermaßen zu verstehen. Da die optisch aktive Platte die Polarisationssebene des Lichtes für verschiedene Frequenzen des Spektrums verschieden stark dreht (siehe für Quarz z. B. die Rotationsdispersion in Abb. 6), so verläßt das Licht die Quarzplatte mit einer Polarisationsrichtung α , die eine Funktion der Frequenz ν des Lichtes ist. Der Einfachheit halber nehmen wir diese Funktion als linear an; messen wir den Winkel α noch von einer passenden Nullrichtung aus, so ist $\alpha = K \cdot \nu$ (K ist dabei eine Konstante).

Betrachten wir irgendeine Stelle unseres Analysators c (Abb. 5 oder 6), deren Polarisationslinien den Winkel β mit der als Nullrichtung gewählten Richtung bilden, so läßt diese Stelle bekanntlich die Intensität

$$I = \cos^2(x - \beta) = \cos^2(K \cdot \nu - \beta)$$

von dem unter dem Winkel $\alpha = K \nu$ polarisierten Licht der Frequenz ν durch. Eine Frequenz ν_1 , für die $K \nu_1 = \beta$ ist, wird mit maximaler Intensität hindurchgehen, eine Frequenz ν_2 mit $K \nu_2 = \beta \pm \pi/2$ wird völlig absorbiert. Je nach der Polarisationsrichtung β der Analysatorstelle wird also eine andere Mischfarbe erscheinen. Da die Analysatorfolie (Abb. 6c) jede Richtung β der Polarisationslinien auf jeweils einem Radiusvektor realisiert, erscheint jede der Mischfarben nur auf einem einzigen Radiusvektor.

Sollen die Mischfarben einen ausgeglichenen Farbenkreis erfüllen, so muß die Konstante K in der obenstehenden Gleichung so bemessen sein, daß $\alpha = K \nu$ etwa 180° überstreicht, wenn ν die Frequenzen vom Rot bis zum Violett durchläuft. Da Tafel I die Drehwinkel für 1 mm Quarz gibt, muß bei Quarz etwa eine

H. Wolter: „Zweidimensionale Farbschlierenverfahren“



a) Glastrichter, annähernd axial gesehen; $0,6 \times$



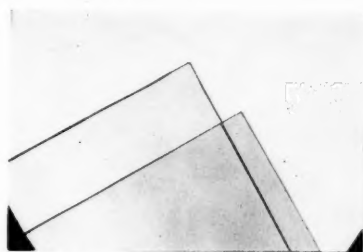
b) Glastrichter, schräg zur Achse gesehen; $0,6 \times$



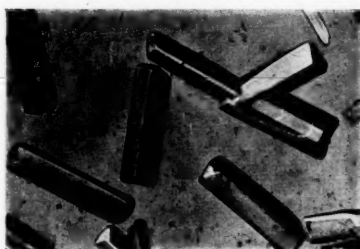
c) Glastrichter, senkrecht zur Achse gesehen; $0,6 \times$



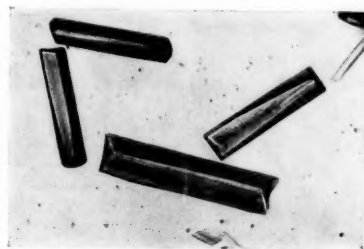
d) Schlechte Glasplatte; $0,6 \times$



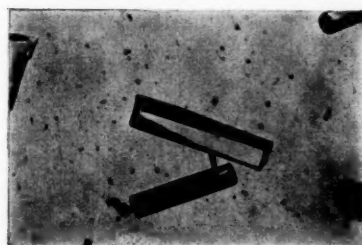
e) Zwei fast planparallele Glasplatten, sich teilweise überdeckend; $1,2 \times$



f) Kaliumsulfatkristalle; $100 \times$



g) Kaliumsulfatkristalle; $100 \times$



h) Kaliumsulfatkristall mit Vizinale; $100 \times$

Aufnahmekamera: Kine-Exakta der Ihagee — Dresden. Film: Agfacolor-Umkehrfilm

JOHANN AMBROSIVS BARTH / VERLAG / LEIPZIG

4 m
da
we

eb
fol
fac
ge
ric
ric
Da
tis
do
gra

ber
Fa
mi
Po

nü
La
un
gra
sch
Ge

mi
ob
sta

an
me
als
da

tei
flä
wä
lie
es
de
do
Gr
mi

4 mm dicke Platte zur Anwendung kommen. Bei einer zu dünnen Platte erscheint das Grün zu stark verweißlicht und das Purpur zu dunkel, bei einer zu dicken werden Gelb und Blau zu weißlich wiedergegeben.

7. Zusammenhang mit einem Polarisationschlierenverfahren

Übrigens ist die im letzten Abschnitt beschriebene Farbsektorplatte natürlich ebenso wirksam, wenn sie um 180° gedreht und mit der „verzerrten“ Polarisationsfolie *c* dem Licht zugewandt wird. Läßt man ferner die Quarzplatte und die einfach linear polarisierende Folie fort, so erhält man ein polarisierendes Schlierengerät, das als Schlierenblende einen Polarisator mit ortsabhängiger Polarisationsrichtung besitzt. Das erzeugte Bild des Objekts kennzeichnet mit der Polarisationsrichtung die Ablenkrichtung und mit dem Polarisationsgrade den Ablenkbetrag. Damit ist ein Verfahren realisiert, wie es wohl H. Schardin bei seinen programmatischen Sätzen auf S. 347 der oben genannten Arbeit vorschwebte. Wie Schardin dort freilich auch bereits betonte, ist das Ergebnis nicht in einem einzigen photographischen Bilde festzuhalten.

Doch haben wir in einer Quarzplatte + Analysator bereits den „Detektor“ beschrieben, um Polarisationsrichtungen in Farbtöne und Polarisationsgrade in Farbsättigungsgrade umzuformen. Würden wir diese Detektorplatte z. B. unmittelbar vor das Bild stellen, so würde sie jene Umwandlung bewirken und die Polarisationszustände auf einem Farbfilm photographierbar machen.

Diese Detektorwirkung ist auch in anderen Fällen, die nicht zu diesem Thema passen, nützlich. So erscheint z. B. ein mit metallischen oder gläsernen Gegenständen erfüllter Laborschrank in leuchtenden Farben, wenn er mit natürlichem weißem Licht beleuchtet und durch jene — an sich ebenfalls farblose — Detektorplatte betrachtet oder photographiert wird. Aus der Farbe und der Farbsättigung, in denen ein Gegenstand dabei erscheint, entnimmt man mit einem Blick, in welcher Weise und in welchem Ausmaß der Gegenstand das Licht polarisiert.

8. Anwendungsbeispiele für das zweidimensionale Farbschlierenverfahren

Das Verfahren kann bei der Untersuchung von Gläsern und Optiken jeder Art mit Nutzen angewandt werden, vor allem aber bei schnell veränderlichen Schlierenobjekten (z. B. ballistischen), die keine zwei Aufnahmen desselben Zustandes gestatten.

Besonders bewährt hat sich das Verfahren bei mikroskopischen Untersuchungen an Kristallen. Schon die Aufnahmen Tafel If-h veranschaulichen das. Wesentlich mehr kann aus Beobachtungen des Kristallwachstums selbst entnommen werden; als Beispiel seien einige Beobachtungen über das Auftreten von Vicinalen kurz dargestellt.

MgSO_4 -Kristalle, die aus warm gesättigter Lösung auskristallisieren, haben teils derbe, teils nadelige Formen. Bei den Nadeln wächst eine Fläche, die Frontfläche, mit einer gegenüber den derben Formen stark erhöhten Geschwindigkeit, während alle Seitenflächen ihre Lage lange Zeit nicht verändern. Wenn bei Vorliegen einer geeigneten Anfangsbedingung ein Nadelwachstum beginnt, so bleibt es wahrscheinlich aus folgenden Gründen erhalten. Die schnell wachsende Fläche der Nadel entzieht der Umgebung die Ionen so weitgehend, daß die Übersättigung dort aufgehoben wird und ein Wachstum der Seitenflächen zunächst aus diesem Grunde unterbleiben muß. Währenddessen finden Lösungsgeossen gemeinsam mit den Wasserdipolen Zeit, die Multipolkräfte der Seitenflächen abzusättigen, so

daß ein Weiterwachsen dort auch nach erneutem Eintreten der Übersättigung nicht erfolgt.

Daß die Lösung in der Tat auch an den Seitenflächen kurze Zeit nach dem Vorüberlaufen der Frontfläche bereits wieder übersättigt ist, beweisen trotz fehlenden Wachstums folgende Beobachtungen.

Wenn die Nadel bei ihrem Wachstum auf eine stark fortschreitende Fläche eines zweiten Kristalls auftritt, so kommen zwar oft die Flächen beider Kristalle in bekannter Weise vor der Berührung zum Stillstand. Trifft aber die Nadel eine „passive“ Seitenfläche einer anderen Nadel, so setzt ein geradezu explosionsartiges Kristallwachstum von der Berührungsstelle aus ein. Von ihr schießen mehrere neue schnellwachsende Kristallnadeln fort, meist einen Winkel von 10 bis 20° mit der getroffenen Fläche bildend; die ältere Nadel, die den ganzen Vorgang verursachte, verdickt sich ebenso plötzlich stark an der Auftreffstelle und wird von hier aus mit einem Kegel überzogen, dessen Winkel gegen die Nadelachse zuerst etwa 26° beträgt und im Laufe der nächsten Sekunden — während die Kegelspitze an der Nadel rückwärts entlangläuft — stetig und monoton abnimmt. Das Auftreten des Kegels beweist die Übersättigung der Lösung in unmittelbarer Nähe der Seitenflächen.

Die Kegelflächen sind in freier Lösung wachsende Vicinalen, wie ihre kontinuierlich sich ändernde Färbung zeigt. Nach etwa 4 Sekunden ist der Kegel so spitz geworden, daß die Flächen der Nadel fast wieder parallel zu den ursprünglichen Nadelflächen verlaufen. Während die Nadel nun nach ihren Seiten mit mäßiger Geschwindigkeit wächst, wird die strenge Lage der Flächen — vielleicht bis auf unmeßbar kleine Restabweichungen — wiederhergestellt. Das Farbschlierenverfahren läßt während des ganzen Vorgangs die Lage der Flächen deutlich, verfolgen.

Erste Versuche mit einem Metallmikroskop zeigten, daß auch hier das Verfahren besonders wirksam ist. Die Farbenplatte wurde zwar bisher nur in den Beleuchtungsteil gesetzt. Ihre Anbringung in der hinteren Brennebene des Objektivs kann aber auch durch eine Zwischenabbildung im Mikroskoptubus umgangen werden; das ist ohne Verschlechterung des Auflösungsvermögens möglich.

Dem Direktor des Instituts, Herrn Professor Dr. Lochte-Holtgreven, dankt der Verfasser für die großzügige Überlassung von Räumen und Geräten.

Kiel, Institut für Experimentalphysik der Universität.

(Bei der Redaktion eingegangen am 27. April 1950.)

Physikalische Begründung eines Farbenkreises und Ansätze zu einer physikalischen Farbenlehre¹⁾

Von Hans Wolter

(Mit 3 Abbildungen)

Inhaltsübersicht

Für die Benutzung der Farben als Kennzeichnungsmittel in der physikalischen Meßtechnik benötigt man einen mathematisch bequemen, physikalisch einfach realisierbaren und genau reproduzierbaren Farbenkreis, dessen Farben bei additiver Mischung stets wieder zu (u. U. noch mit Grau vermischten) Farben desselben Farbenkreises führen. Farben werden dabei als gleich bezeichnet, wenn ihre spektrale Zusammensetzung identisch ist.

Die genannten Forderungen werden von den „Normalfarben“ erfüllt. Ihre Eigenschaften sind in der Zusammenfassung am Schluß der Arbeit aufgeführt.

§ 1. Einleitung

Die große Mannigfaltigkeit der dem menschlichen Auge unterscheidbaren Farben erhält bekanntlich eine übersichtliche Ordnung durch den Farbenkreis, d. h. jenes eindimensionale zyklische Kontinuum von Farben, das man erhält, wenn man die Spektralfarben durch die Purpurgemische (Gemische eines gerade noch sichtbaren Violett mit gerade noch sichtbarem Rot) ergänzt. Denn jede Farbe ist genau einer (mit Grau vermischten) Farbe des Farbenkreises physiologisch äquivalent.

Insbesondere ist die additive Mischung von Farben des Farbenkreises stets einer Mischung einer bestimmten Farbe des Farbenkreises mit Grau äquivalent; diese letzte Eigenschaft im Zusammenhang mit der zyklischen Struktur macht die Farben des Farbenkreises geeignet als Kennzeichnungsmittel in der physikalischen Meßtechnik, und zwar allgemein überall dort, wo den Elementen einer zweidimensionalen oder dreidimensionalen Mannigfaltigkeit einfach beobachtbare Kennzeichen eineindeutig und stetig zugeordnet werden sollen. Ein Beispiel hierfür ist das zweidimensionale Farbschlierenverfahren²⁾, das die Ablenkungen des Lichtes durch ein Schlierenobjekt nach Ablenkrichtung und Ablenkbetrag erkennen läßt. Dabei erhält das Bild des Objekts an jeder Stelle eine Färbung, deren Farbton die Ablenkrichtung und deren Sättigung den Ablenkbetrag kennzeichnet. Bei dreidimensionalen Mannigfaltigkeiten könnte noch die Intensität als drittes Kennzeichen dienen.

Bei diesem Farbschlierenverfahren und ebenso bei anderen geplanten Anwendungen kommt nun die resultierende Farbe im Bilde durch additive Mischung vieler

¹⁾ Vorgetragen auf der Tagung der Deutschen Physikalischen Gesellschaft in Münster am 15. April 1950.

²⁾ Ann. Physik (6) 8, 1 (1950).

Farben des Farbenkreises zustande. Die Mischfarbe ist zwar selbstverständlich einer grauvermischten Farbe des Farbenkreises physiologisch äquivalent; doch ist diese Äquivalenz abhängig von dem Auge des Beobachters, da die beiden zu vergleichenden Farben zwangsläufig sehr verschiedene spektrale Zusammensetzung haben. Besonders unangenehm ist, daß die Aussagen der Beobachter über gleiches Aussehen von spektral verschieden zusammengesetzten Farben auch zeitlich sehr schwanken. Das zeigte sich, als Verfasser eine Neuvermessung der König-Dieterici-Kurven in Angriff genommen hatte. Diese entsprachen dem Auge des Verfassers zwar nicht; aber die Variationsbreite erwies sich schließlich als so groß, daß eine Berichtigung der KD-Kurven für die physikalische Anwendung der Farben ohnehin nicht ausgereicht hätte. So blieb nur noch übrig, die physikalischen Anwendungen der Farben ganz von Gesichtspunkten der physiologischen Äquivalenz zu befreien.

Der erste Schritt auf diesem Wege besteht in der Schaffung eines Farbenkreises, dessen Farben miteinander additiv gemischt stets wieder zu einer (u. U. grauvermischten) Farbe desselben Kreises im Sinne der physikalischen Gleichheit führen. Die Frage, ob es einen solchen Farbenkreis — wir nennen ihn einen „Normalfarbenkreis“ — überhaupt gibt, wird in dieser Arbeit bejaht werden; wir werden die spektrale Zusammensetzung der „Normalfarben“ mathematisch beschreiben, Methoden zur experimentellen Realisierung aufweisen und die Frage nach der Eindeutigkeit der Lösung beantworten.

§ 2. Grundlegende Definitionen

Df. 1: Unter einer Farbe verstehen wir eine stückweise stetige, stückweise monotone Funktion $f(v) \geq 0$ der Grundvariablen v . Stets sei

$$f(v_0) = \frac{1}{2} \left(\lim_{v \rightarrow v_0 + 0} f(v) + \lim_{v \rightarrow v_0 - 0} f(v) \right).$$

Die Grundvariable v ist selbst eine beliebige eigentlich monotone stetige Funktion der Wellenlänge λ . Ist $f(v)$ eine Konstante, so sprechen wir von einem „Grau“; $f(v) \equiv 1$ heiße „Weiß“, $f(v) \equiv 0$ heiße „Schwarz“.

Der Leser möge sich um der Anschaulichkeit willen v als Frequenz denken, bis im § 7 näher auf die Vorteile gewisser $v(\lambda)$ eingegangen wird.

Einer Körperoberfläche schreiben wir als Farbe ihre Remissionsfunktion zu, einem Filter das Verhältnis der durchgelassenen zur auffallenden Intensität als Funktion von v .

Falls die so definierten Funktionen von der spektralen Verteilung des einfallenden Lichtes abhängen (z. B. bei fluoreszierenden Stoffen), ist das einfallende Licht als energiegleiches Spektrum bezüglich v als der Abszisse anzusetzen (d. h. die in ein Intervall $(v_1; v_2)$ fallende Intensität sei proportional der Intervalllänge $|v_2 - v_1|$).

Die von der einfallenden Strahlung unabhängig emittierte Strahlung der betrachteten Körper ist dabei als verschwindend klein gegenüber der reflektierten oder hindurchgelassenen Strahlung vorausgesetzt.

Für die oben definierten Farben gilt in den weitaus meisten Fällen $f(v) \leq 1$. Doch wollen wir das nicht voraussetzen. Farben, bei denen es ein v_1 gibt, so daß $f(v_1) > 1$ ist, nennen wir überhöhte Farben.

Von der Farbe einer Strahlung wollen wir nur im Vergleich mit einer anderen Strahlung sprechen und dann das Intensitätsverhältnis als Funktion von v darunter verstehen.

Df. 2: Farben $f_1(v)$ und $f_2(v)$ heißen gleich, wenn $f_1(v) = f_2(v)$ für jedes v .

Df. 3: Ist $f(v; \mu)$ eine kontinuierliche Menge von Farben, deren jede durch eine reelle Zahl μ gekennzeichnet ist, und ist $\varrho(\mu) \geq 0$ eine Funktion mit $\int \varrho(\mu) d\mu = 1$, so versteht man unter

$$\int \varrho(\mu) f(v; \mu) d\mu$$

die additive Mischung der Farben $f(v; \mu)$ zur Gewichtsfunktion $\varrho(\mu)^{2)}$.

Bezeichnet man mit $\delta(x)$ die Sommerfeldsche Zackenfunktion und wählt speziell

$$\varrho(\mu) = \frac{1}{2} \{ \delta(\mu - \mu_1) + \delta(\mu - \mu_2) \},$$

so besagt die Df. 3 speziell

Df. 3a: Unter der additiven Mischung zweier Farben $f_1(v)$ und $f_2(v)$ versteht man die Farbe $\frac{1}{2} \{ f_1(v) + f_2(v) \}$.

Auch ohne Verwendung der uneigentlichen Funktion $\delta(x)$ können die additive Mischung diskreter und kontinuierlicher Farbenmengen einheitlich formuliert werden, wenn man statt des Riemannschen Integrals das Stieltjesintegral verwendet.

Wenn wir mit den Definitionen 3 und 3a von denen Littmanns⁴⁾ und anderer abweichen, so geschieht das mit Rücksicht auf drei Vorteile in der Praxis. Die Mischung zweier Pigmentfarben mit Remissionsfunktionen f_1 und f_2 zu gleichen Teilen führt unmittelbar zu $\frac{1}{2} (f_1 + f_2)$, sofern nur die einzelnen Pigmentteilchen optisch dick und groß gegen die Wellenlänge sind. Für die Mischung der Filterfarben durch Übereinanderprojizieren zweier gefilterter Strahlenbündel ist es naheliegend, vor Passieren beider Filter eine Trennung eines Strahles (durch halb durchlässig versilberten Spiegel oder dergleichen) vorauszusetzen und die nach der Vereinigung resultierende Intensität auf die Intensität vor der Trennung zu beziehen; das Ergebnis dieser Farbenmischung entspricht unserer Definition. Vor allem aber ist unsere Definition ersichtlich besonders einfach auf die für das Farbschlierenverfahren u. ä. wichtige kontinuierliche Farbenmischung anzuwenden.

Df. 4: Unter multiplikativer Mischung zweier Farben $f_1(v)$ und $f_2(v)$ versteht man die Farbe $f_1(v) \cdot f_2(v)$.

Oft wurde diese Farbenmischung als subtraktive bezeichnet. Wollte man auch die (freilich in der Praxis kaum interessierende) multiplikative Mischung kontinuierlicher Farbenmengen erfassen, so wäre eine Definition durch $p_0(v) = \exp(\int \varrho(\mu) \ln f(v, \mu) d\mu)$ zweckmäßig; würde man wieder $\int \varrho(\mu) d\mu = 1$ voraussetzen, so erhielte man für die multiplikative Mischung zweier Farben $\sqrt{f_1(v) f_2(v)}$. Doch empfiehlt sich das für die Praxis nicht; denn zwei hintereinandergesetzte Filter oder ein mit zwei Lasurfarben übereinander gestrichenes Papier lassen sich mit unserer Definition 4 bequemer erfassen.

Ersichtlich führt beliebig kombinierte additive und multiplikative Mischung nichtüberhöhter Farben stets wieder zu einer nichtüberhöhten Farbe.

§ 3. Mathematische Formulierung der Forderungen an einen physikalischen Farbenkreis und an einen Normalfarbenkreis

Eine jede Farbe können wir uns veranschaulichen, indem wir der Abb. 1 entsprechend eine Schablone über das Spektrum decken und das von der Schablone

²⁾ Alle Integrale in dieser Arbeit sind bestimmte Integrale. Sind keine Grenzen angegeben, so ist die Integration über den gesamten Definitionsbereich der Integrationsvariablen zu erstrecken.

⁴⁾ Littmann, Optik 5, 180 (1949).

hindurchgelassene Licht gemischt denken. Die Schablonenkante gibt dann die Farbe $f(v)$.

Unter einem physikalischen Farbenkreis verstehen wir eine Teilmenge aus der Menge aller Farben, die in folgender Weise erzeugt werden kann. Eine periodische Schablone werde längs der v -Achse verschoben; jede Schablonenstellung führt zu einer Farbe des Farbenkreises. Ist x die in v -Richtung gemessene Koordinate auf dem Schablonenblech und gibt $F(x)$ die Schablonenkante, so wird die Farbe $F(v-\mu)$ von der Schablone durchgelassen. Dabei gibt μ die Stelle der v -Achse, an der sich der schablonenfeste Koordinatenanfang der x -Achse befindet. Die sowohl in v als auch in μ periodische Funktion $F(v-\mu)$ repräsentiert für jedes μ eine Farbe des Farbenkreises.

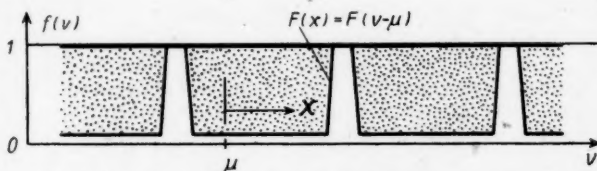


Abb. 1. Zur Erzeugung eines physikalischen Farbenkreises durch Verschieben einer periodischen Schablone im Spektrum.

Df. 5: $F(v-\mu)$ bezeichnen wir als physikalischen Farbenkreis, wenn

a) $F(x)$ stückweise stetig und stückweise monoton ist, ferner an jedem Argument einen rechtsseitigen und einen linksseitigen Limes besitzt, deren arithmetisches Mittel gleich dem Funktionswert ist;

b) $0 \leq F(x) \leq 1$;

c) $F(x+g) = F(x)$; (Periodizitätsforderung; g bleibt verfügbar).

Df. 6: Einen physikalischen Farbenkreis $F(v-\mu)$ bezeichnen wir als einen „Normalfarbenkreis“, wenn

d) es zu jeder Funktion $\varrho(\mu)$ mit $\varrho(\mu) \geq 0$; $\int \varrho(\mu) d\mu = 1$ ein $A_0 \geq 0$, ein $B_0 \geq 0$ und ein reelles μ_0 so gibt, daß

$$\int \varrho(\mu) F(v-\mu) d\mu = A_0 + B_0 \cdot F(v-\mu_0). \quad (1)$$

(Additivitätsforderung; sie verlangt entsprechend der in der Einleitung formulierten Aufgabe, daß additive Mischung von Farben des Normalfarbenkreises stets zu einer (u. U. grauverschmutzten) Farbe desselben Farbenkreises führt.)

Alle durch additive Mischung aus Farben des Normalfarbenkreises entstehenden Farben nennen wir dann „Normalfarben“.

§ 4. Trigonometrische Funktionen als einzige Lösungen des Problems

A) Wir nehmen in diesem Abschnitt A an, daß es mindestens eine Funktion $F(x)$ gibt, die den Bedingungen a bis d genügt; dann kann diese wegen der Voraussetzungen a und c in eine Fourierreihe entwickelt werden

$$F(x) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i \frac{x}{g} n}, \quad (2)$$

die für alle x konvergiert und die Funktion darstellt. Einsetzen in die Bedingung d führt wegen der gliedweisen Integrierbarkeit von Fourierreihen zu

$$\sum_{n=-\infty}^{\infty} a_n e^{2\pi i \frac{v}{g} n} \cdot \int \varrho(\mu) e^{-2\pi i \frac{\mu}{g} n} = A_e + B_e \cdot \sum_{n=-\infty}^{\infty} a_n e^{2\pi i \frac{v-\mu}{g} n} \quad (3)$$

und wegen der Eindeutigkeit der Fourierreihenentwicklung durch Koeffizientenvergleich zu

$$a_n \cdot \int \varrho(\mu) e^{-2\pi i \frac{\mu}{g} n} d\mu = a_n B_e e^{-2\pi i \frac{\mu_0}{g} n} \quad \text{für } n \neq 0. \quad (4)$$

Sind n und m zwei von Null verschiedene Indizes mit $m \geq n$, für die $a_n \neq 0$ und $a_m \neq 0$, so muß also

$$\left| \int \varrho(\mu) e^{-2\pi i \frac{\mu}{g} n} d\mu \right| = \left| \int \varrho(\mu) e^{-2\pi i \frac{\mu}{g} m} d\mu \right| \quad (5)$$

sein für jedes $\varrho(\mu)$ mit $\int \varrho(\mu) d\mu = 1$, also auch speziell für

$$\varrho(\mu) = \frac{1}{2} \{ \delta(\mu - \mu_1) + \delta(\mu - \mu_2) \}. \quad (6)$$

Dafür besagt Gl. (5)

$$\left| \cos\left(\frac{\pi}{g} |n| (\mu_1 - \mu_2)\right) \right| = \left| \cos\left(\frac{\pi}{g} |m| (\mu_1 - \mu_2)\right) \right|. \quad (7)$$

Wählen wir z. B. speziell $|\mu_1 - \mu_2|$ so, daß $\frac{\pi}{g} |m| |\mu_1 - \mu_2| = \frac{\pi}{4}$ ist, so ist (7) nur zu erfüllen, wenn

$$|n| = |m| \quad (8)$$

ist. Die Fourierreihe muß sich also zwangsläufig reduzieren auf eine Funktion

$$F(x) = a_0 + a_n e^{2\pi i \frac{x}{g} n} + a_{-n} e^{-2\pi i \frac{x}{g} n}, \quad \text{also auf}$$

$$F(x) = \text{Konst.} + \text{Konst.} \cos\left(2\pi \frac{n}{g} (x - \text{Konst.})\right) \quad \text{oder} \quad (9)$$

$$F(x) = A + B \cdot \cos^2(\pi a x'),$$

wenn x' sich von x durch eine Konstante unterscheidet. Wählen wir o. B. d. A. den Anfangspunkt der Koordinate x passend, d. h. die Marke auf der Schablone, mit deren Hilfe wir die Lage der Schablone im Spektrum kennzeichnen, so muß also $F(v - \mu)$ in der Form

$$F(v - \mu) = A + B \cos^2(\pi a (v - \mu)); \quad A \geq 0; \quad B \geq 0; \quad A + B \leq 1 \quad (10)$$

darstellbar sein, sofern $F(v - \mu)$ ein Normalfarbenkreis sein soll. Die Größe a , d. h. die Periodenlänge $g = \frac{n}{a}$, ist mit unseren bisherigen physikalischen For-

derungen natürlich nicht festgelegt; wir werden sie später nach der Breite des sichtbaren Spektrums, also nach einem physiologischen Gesichtspunkt, bemessen.

B) Daß die Farben (10) tatsächlich Normalfarben sind (d. h. den Bedingungen *a* bis *d* der Definitionen 5 und 6 genügen), ist bezüglich *a* bis *c* evident; für *d* folgt es unmittelbar aus dem Additionstheorem der trigonometrischen Funktionen; denn es ist, wie eine elementare Rechnung zeigt, für jedes $\varrho(\mu) \geq 0$ mit $\int \varrho(\mu) d\mu = 1$

$$\begin{aligned} \int \varrho(\mu) \cos^2(\pi a(v-\mu)) d\mu &= \frac{1}{2} \{1 + C_e \cos(2\pi a v) + S_e \sin(2\pi a v)\} \\ &= \frac{1-B_e}{2} + B_e \cos^2(\pi a(v-\mu_0)) \end{aligned} \quad (11)$$

mit

$$\begin{aligned} C_e &= \int \varrho(\mu) \cos(2\pi a \mu) d\mu; \quad S_e = \int \varrho(\mu) \sin(2\pi a \mu) d\mu \\ B_e &= \sqrt{C_e^2 + S_e^2} \end{aligned} \quad (12)$$

$$\cos(2\pi a \mu_0) = \frac{C_e}{B_e}; \quad \sin(2\pi a \mu_0) = \frac{S_e}{B_e}.$$

§ 5. Proportionalklassen, reduzierte Normalfarben, geometrische Veranschaulichung der Normalfarben

Farben, die auseinander lediglich durch Vermischen mit Schwarz hervorgehen, sind so verwandt, daß man sich oft begnügen kann, aus ihnen jeweils einen Repräsentanten zu betrachten. Wir definieren daher:

Df. 7: Zwei Farben $f_1(v)$ und $f_2(v)$ heißen einander proportional, in Zeichen $f_1(v) \sim f_2(v)$, wenn mindestens eine der folgenden Bedingungen erfüllt ist:

- Es gibt ein c_1 , so daß $c_1 f_1(v) = f_2(v)$ für alle v ;
- Es gibt ein c_2 , so daß $c_2 f_2(v) = f_1(v)$ für alle v .

Df. 8: Unter einer Proportionalklasse verstehen wir die Menge aller Farben, die zu einer festen nicht schwarzen Farbe proportional sind.

Die Farbe Schwarz ist jeder Farbe proportional, sie ist also Element jeder Proportionalklasse. Jede Farbe ist sich selbst proportional; jede nicht schwarze Farbe gehört daher genau einer Proportionalklasse an.

Df. 9: Die Farben $\cos^2(\pi a(v-\mu))$ nennen wir gesättigte Normalfarben, ihre Gesamtheit einen gesättigten Normalfarbensenkreis. Alle durch additive Mischung aus ihm hervorgehenden Farben (11) nennen wir reduzierte Normalfarben.

Zu jeder Farbe (10) gibt es genau eine ihr proportionale reduzierte Normalfarbe der Gestalt

$$\frac{1-B}{2} + B \cos^2(\pi a(v-\mu)); \quad 0 \leq B \leq 1; \quad (13)$$

das ist eine gesättigte Normalfarbe für $B = 1$; $B = 0$ bedeutet reduziertes Grau. Jede Farbe (10) läßt sich in der Form schreiben

$$2I \left\{ \frac{1-B}{2} + B \cos^2(\pi a(v-\mu)); \quad 0 \leq B \leq 1. \right. \quad (14)$$

Df. 10: Bei einer Farbe (14) bezeichnen wir μ als Farbton. B als Sättigung und I als Mittelwert.

Reduzierte Normalfarben sind Normalfarben mit dem Mittelwert 0,5.

Alle Normalfarben nach (14) bilden eine dreidimensionale Mannigfaltigkeit. Doch wollen wir uns zur Vereinfachung nur für die zweidimensionale Mannigfaltigkeit der Proportionalklassen von Normalfarben interessieren; da jede Proportionalklasse von Normalfarben genau eine reduzierte Normalfarbe enthält, wählen wir diese als Repräsentanten der Proportionalklasse und befassen uns nur mit diesen. Sie bilden nach (13) eine zweidimensionale Mannigfaltigkeit, die wir durch die Festsetzung $r = B$, $\varphi = 2\pi\alpha\mu$ den Punkten des Kreisesinnern ($r \leq 1$) mit den Polarkoordinaten r, φ eineindeutig zuordnen können. (Siehe Abb. 2!)

Den Rand des Kreises nehmen die Normalfarben mit der Sättigung $B = 1$ ein, sie bilden den gesättigten Normalfarbenkreis. Die Mitte $B = 0$ stellt das reduzierte Grau ($f(v) = \frac{1}{2}$) dar.

Der Wert der geometrischen Zuordnung liegt in der Übersichtlichkeit der geometrischen Konstruktion, die den analytischen Operationen (11), (12) entspricht.

Der in Gl. (11) ausgeführten additiven Mischung gesättigter Normalfarben entspricht nämlich geometrisch einfach die Schwerpunktsbildung; denn zwischen den Cartesischen Koordinaten ξ_e, η_e der resultierenden Farbe und denen der Ausgangsfarben $\xi(\mu) = \cos(2\pi\alpha\mu)$, $\eta(\mu) = \sin(2\pi\alpha\mu)$ bestehen nach Gl. (11) und (12) offenbar die Beziehungen

$$\begin{aligned}\xi_e &= B_e \cos(2\pi\alpha\mu_e) = C_e = \int \varrho(\mu) \cos(2\pi\alpha\mu) d\mu = \int \varrho(\mu) \xi(\mu) d\mu; \\ \eta_e &= B_e \sin(2\pi\alpha\mu_e) = S_e = \int \varrho(\mu) \sin(2\pi\alpha\mu) d\mu = \int \varrho(\mu) \eta(\mu) d\mu.\end{aligned}\quad (15)$$

Auch die additive Mischung von nicht gesättigten Normalfarben führt, wie ebenso zu zeigen, stets zu dem Schwerpunkt der Ausgangsfarben.

Ist insbesondere die Gewichtsfunktion $\varrho(\mu)$ symmetrisch zu einem μ_0 , so ist das bei der Mischung resultierende $\mu_e = \mu_0$. Das hat eine wichtige Konsequenz für das oben genannte Farbschlierenverfahren. Dort kam die Mischung so zustande, daß ein Lichtbündel kreisförmigen Querschnitts eine Farbsektorplatte durchsetzte. Ist diese Farbsektorplatte nun eine ebene Platte, die an den Punkten mit Polarkoordinaten R, α mit den gesättigten Normalfarben $\cos^2(\pi\alpha(v-\mu)) = \cos^2\left(\pi\alpha v - \frac{\alpha}{2}\right)$ angefärbt ist (eine einfache experimentelle Realisierung wurde übrigens bereits in der Veröffentlichung über das Farbschlierenverfahren beschrieben), so erfährt ein exzentrisch hindurchtretendes Lichtbündel kreisförmigen Querschnitts nach der Strahlenvereinigung eine Farbmischung, deren Gewichtsfunktion

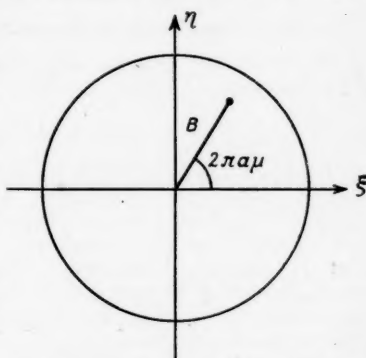


Abb. 2. Zuordnung der reduzierten Normalfarben

$\frac{1}{2}(1 - B) + B \cos^2(\pi\alpha(v - \mu))$
zu den Punkten $|\xi^2 + \eta^2| \leq 1$ einer Ebene

durch die Lage des Kreisbündels gegeben ist und jedenfalls symmetrisch zu $\mu_0 = \alpha_0/2\pi a$ ist, wenn der Kreisquerschnitt des Bündels auf der Farbsektorplatte symmetrisch zu dem Radius in Richtung α_0 liegt. Das zu der Farbe gehörige $\mu_0 = \mu_0$ ist also nur von der Richtung, in die der Kreis aus der zentralen Lage abgelenkt ist, abhängig und nicht von dem Ablenkbetrag. Folglich kennzeichnet das Farbschlierenverfahren, wenn es mit Normalfarben durchgeführt wird, die Ablenkrichtung durch einen eindeutigen Farbton; der Ablenkbetrag wird durch die Sättigung gekennzeichnet.

Diese Zusammenhänge gelten bei Verwendung von Normalfarben im Sinne der strengen Farbgleichheit unserer Df. 2, während Entsprechendes mit anderen Farbenkreisen höchstens im Sinne der Farbenäquivalenz erreichbar ist.

§ 6. Einordnung der Normalfarben in die Äquivalenzklassen⁵⁾

Df. 11: Ist $b(v) \geq 0$ eine stückweise stetige und stückweise monotone Funktion mit existierendem $\int b(v) dv$, so heißt

$$x_b = \int b(v) f(v) dv \quad (16)$$

die Komponente der Farbe $f(v)$ nach der Funktion $b(v)$.

Df. 12: Ist \mathfrak{M} eine beliebige Menge (wir nennen sie „Dimensionsmenge“) und liegt für jedes Element i dieser Menge eine Funktion $b_i(v)$, die die Voraussetzungen der Df. 11 erfüllt, vor, sind ferner alle $b_i(v)$ linear unabhängig voneinander, so nennen wir die Gesamtheit der $b_i(v)$ eine Basis von der Dimension $|\mathfrak{M}|$. Mit $|\mathfrak{M}|$ sei darin die Kardinalzahl der Dimensionsmenge \mathfrak{M} bezeichnet. Die nach Df. 11 gebildeten Komponenten x_i einer Farbe $f(v)$ nach den $b_i(v)$ einer Basis nennen wir die Koordinaten der Farbe nach der Basis.

Df. 13: Haben die Koordinaten einer Farbe nach einer Basis eine von der Reihenfolge der Summanden unabhängige Summe

$$\sum_{i \in \mathfrak{M}} x_i = s \neq 0, \quad (17)$$

so heißen die

$$x'_i = \frac{x_i}{s} \quad (18)$$

die normierten Koordinaten der Farbe nach der vorgelegten Basis. Ersichtlich existiert dann die Summe der x'_i und ist gleich 1.

Df. 14: Farben $f_1(v)$ und $f_2(v)$ heißen äquivalent bezüglich der Basis $b_i(v)$, in Zeichen

$$f_1(v) \stackrel{=}{b_i} f_2(v),$$

wenn $f_1(v)$ und $f_2(v)$ identische Koordinaten nach der Basis $b_i(v)$ haben. Zwei Farben heißen äquivalentproportional, in Zeichen

$$f_1(v) \stackrel{\sim}{b_i} f_2(v),$$

⁵⁾ Die Lektüre dieses Paragraphen ist zum Verständnis der folgenden nicht erforderlich.

wenn sie gleiche normierte Koordinaten bezüglich der Basis haben. Farben, deren Koordinaten gleich ihren normierten Koordinaten sind, heißen normierte Farben.

Df. 15: Alle einander äquivalenten Farben bilden eine Äquivalenzklasse; alle einander äquivalentproportionalen bilden eine Äquivalentproportionalklasse.

Df. 16: Zwischen Äquivalenzklassen definieren wir eine additive Mischung, indem wir die Df. 3 und 3a auf ein beliebiges Element (einen „Repräsentanten“) der Äquivalenzklassen anwenden. Man kann leicht zeigen, daß diese Df. stets zu einem eindeutigen (d. h. von der Wahl des Repräsentanten unabhängigen) Ergebnis führt.

Df. 17: Zwischen Äquivalentproportionalklassen definieren wir eine additive Mischung, indem wir die Df. 3 und 3a jeweils auf eine normierte Farbe als Repräsentanten der Äquivalentproportionalklasse anwenden.

Df. 18: Gibt es zu einer Basis $b_i(v)$ und einer zweiten Basis $b_i^*(v)$ Zahlen a_{ik} , so daß

$$b_i^*(v) = \sum_{k \in \mathfrak{M}} a_{ik} b_k(v); \quad i \in \mathfrak{M}, \quad \text{für alle } v, \quad (19)$$

so sagt man, beide Basen seien miteinander verwandt; Gl. (19) heißen die Transformationsgleichungen einer „homogenen linearen Koordinatentransformation“.

Wegen der Linearität des Integraloperators in (16) bleibt bei homogener linearer Koordinatentransformation der Elementenbestand einer Äquivalenzklasse (und ebenso einer Äquivalentproportionalklasse) stets erhalten; m. a. W.: Farben, die bezüglich einer Basis äquivalent sind, sind es auch bezüglich einer verwandten Basis; Entsprechendes gilt für die Äquivalentproportionalität. Bei Übergang von einer Basis zu einer zur ersten nicht verwandten kann dagegen die Äquivalenz zweier Farben verlorengehen.

Die Äquivalenzklassen zu einer Basis von der Dimension $|\mathfrak{M}|$ kann man geometrisch veranschaulichen, indem man der Äquivalenzklasse mit den Koordinaten x_i den Punkt des $|\mathfrak{M}|$ -dimensionalen linearen Raumes R mit den Koordinaten x_i zuordnet. Dem Übergang von einer Basis zu einer dazu verwandten entspricht dann eine lineare homogene Transformation des Koordinatensystems. Um der einfacheren Sprechweise willen werden wir Punkte dieses Raumes geradezu Äquivalenzklassen nennen.

Alle auf dem gleichen Ursprungsstrahl liegenden Äquivalenzklassen enthalten Farben, die zu einer Menge vereinigt, gerade eine Äquivalentproportionalklasse bilden. Da für die normierten Koordinaten x_i' nach Gl. (18) stets

$$\sum_{i \in \mathfrak{M}} x_i' = 1 \quad (20)$$

gilt, liegen die zu normierten Farben gehörigen Äquivalenzklassen stets in dem $(|\mathfrak{M}| - 1)$ -dimensionalen linearen Unterraum R_p , der durch $\sum x_i = 1$ definiert ist. Jeder Äquivalentproportionalklasse können wir daher eindeutig einen Punkt dieses Unterraumes zuordnen; umgekehrt entspricht jedem Punkt dieses Unterraumes höchstens eine Äquivalentproportionalklasse. Da definitionsgemäß alle Komponenten nicht negativ sind, wird für unsere geometrische Veranschaulichung nur der Teil des $|\mathfrak{M}|$ -dimensionalen Raumes R und ebenso des $(|\mathfrak{M}| - 1)$ -dimensionalen Unterraumes R_p benötigt, dessen Punkte nichtnegative Koordinaten haben.

Werden Äquivalenzklassen $x_i(\mu)$ nach Df. 16 (Repräsentanten seien $f(v; \mu)$) miteinander gemischt, so entspricht dem geometrisch die Schwerpunktsbildung; denn die resultierende Äquivalenzklasse hat die Koordinaten

$$\bar{x}_i = \int b_i(v) \int \varrho(\mu) f(v; \mu) d\mu dv = \int \varrho(\mu) \int f(v; \mu) b_i(v) dv d\mu = \int \varrho(\mu) x_i(\mu) d\mu.$$

Werden Äquivalenzklassen normierter Farben miteinander additiv gemischt, so ist das Ergebnis (Schwerpunkt) wegen der Linearität des Unterraumes (20) wieder eine Äquivalenzklasse normierter Farben. Wegen der Df. 17 bleibt dieser Satz richtig, wenn statt Äquivalenzklasse normierter Farben durchweg Äquivalentproportionalklasse gesagt wird.

Die gesättigte Normalfarbe $\cos^2(\pi a(v - \mu))$ hat bzgl. der Basis $b_i(v)$ die Koordinaten

$$x_i(\mu) = \frac{1}{2} + \frac{1}{2} \cos(2\pi a \mu) \cdot \int b_i(v) \cos(2\pi a v) dv + \frac{1}{2} \sin(2\pi a \mu) \cdot \int b_i(v) \sin(2\pi a v) dv. \quad (21)$$

Das folgt aus dem Additionstheorem der trigonometrischen Funktionen in derselben Weise wie Gl. (11). Die $|\mathfrak{M}|$ Gleichungen (21) bedeuten eine Kurve im Raume R in Parameterdarstellung mit dem Parameter μ . Falls $|\mathfrak{M}| \geq 3$ ist, können aus dreien dieser Gleichungen $\sin(2\pi a \mu)$ und $\cos(2\pi a \mu)$ durch rein lineare Operationen eliminiert werden. Das Ergebnis dieser Elimination ist eine die x_i linear enthaltende Gleichung, die einen $(|\mathfrak{M}| - 1)$ -dimensionalen linearen Unterraum R_1 definiert; ihm gehört daher die gesamte Kurve (21) an. Da die ungesättigten reduzierten Normalfarben durch additive Mischung aus den gesättigten hervorgehen, die zugehörigen Äquivalenzklassen also Schwerpunkte irgendwelcher Punkte unserer Kurve (21) sind, gehören also alle reduzierten Normalfarben, die zu einem $a \cdot v(\lambda)$ gehören, dem genannten $(|\mathfrak{M}| - 1)$ -dimensionalen linearen Unterraum R_1 an.

Rechnet man ferner aus irgend zweien der Gleichungen (21) $\sin(2\pi a \mu)$ und $\cos(2\pi a \mu)$ aus, quadriert und addiert, so erhält man eine Gleichung, auf deren einer Seite 1 und auf deren anderer Seite eine quadratische Form der x_i steht; diese Gleichung definiert einen $(|\mathfrak{M}| - 1)$ -dimensionalen Unterraum zweiten Grades R_2 , dem die Äquivalenzklassen gesättigter Normalfarben ebenfalls angehören müssen. Die Kurve (21) ist also als „Schnittkurve“ zweier $(|\mathfrak{M}| - 1)$ -dimensionaler Unterräume, eines linearen R_1 und eines vom 2. Grade, selbst eine Kurve 2. Grades.

Projizieren wir diese vom Nullpunkt aus auf den linearen Unterraum R_p ($\sum x_i = 1$), so erhalten wir wieder eine Kurve 2. Grades als den geometrischen Ort aller Proportionalklassen der gesättigten Normalfarben.

Diese Verhältnisse werden anschaulich vor allem, wenn wir die Dimension $|\mathfrak{M}| = 3$ wählen, wie das bei der verbreiteten Farbenlehre geschieht. Der Raum R ist dann ein dreidimensionaler linearer Raum. Der lineare Unterraum R_p für die Proportionalklassen ist die Ebene

$$x_1 + x_2 + x_3 = 1,$$

die alle drei Koordinatenachsen in dem Abstände 1 vom Ursprung schneidet.

Die Äquivalenzklassen der gesättigten Normalfarben $\cos^2(\pi a(v - \mu))$ erfüllen eine Kurve (21), die als Schnitt einer Ebene R_1 mit einer Fläche 2. Grades R_2 eine Kurve 2. Grades sein muß. Ihre Projektion auf die Ebene R_p vom Ursprung

als Projektionszentrum aus ist dann zwangsläufig ebenfalls eine Kurve 2. Grades. Da diese überdies ganz in dem Raumteil mit nichtnegativen Koordinaten bleiben muß, kann sie das von den Koordinatenebenen begrenzte gleichseitige Dreieck der Ebene R_p

$$x_1 + x_2 + x_3 = 1; \quad x_1 \geq 0; \quad x_2 \geq 0; \quad x_3 \geq 0 \quad (22)$$

nicht verlassen, kann also nur eine Ellipse sein.

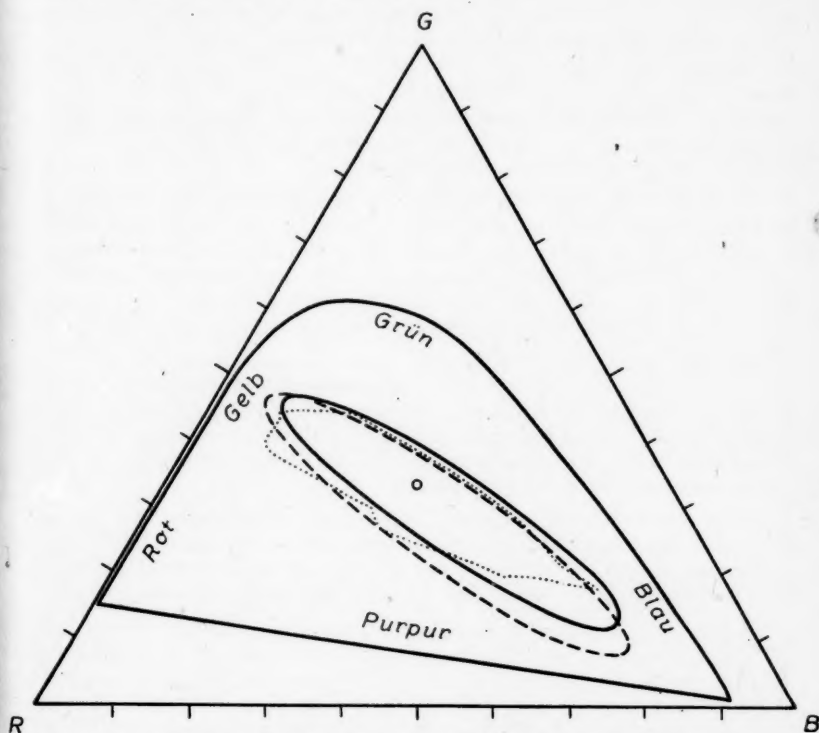


Abb. 3. Darstellung der Äquivalentproportionalklassen im Farbdreieck. Als Basis dienen die 3 König-Dieterici-Kurven. Parabelartiger Bogen: Spektralfarben, geschlossen durch die Purpurgerade. Punktierte Kurve: Ostwaldscher Farbenkreis *na*. Ausgezogene Ellipse: Normalfarben zu ν als Frequenz und Periodenlänge $g = 300 \cdot 10^{12}$ Hz. Gestrichelte Ellipse: Dasselbe für $g = 400 \cdot 10^{12}$ Hz

Abb. 3 zeigt ein solches Dreieck (22)⁶⁾. Die ausgezogene Ellipse gibt die Äquivalentproportionalklassen der gesättigten Normalfarben zu einer frequenzproportionalen Grundfunktion; und zwar wurde $\nu(\lambda)$ als Frequenz $\frac{c}{\lambda}$ und $g = \frac{1}{a} =$

⁶⁾ Wir wählen dabei die gleichseitige Form, um die zu Ostwaldschen Farben gehörenden Punkte unmittelbar der Arbeit K. W. F. Kohlrausch, Physik. Z. 21, 403 (1920) entnehmen zu können.

300 · 10¹² Hz gewählt. Als Basis $b_i(\nu)$ dienten die auf Frequenzabszisse umgerechneten König-Dieterichschen „Grundempfindungskurven“.

Die gestrichelte Ellipse gibt einen Normalfarbenkreis mit $g = \frac{1}{a} = 400 \cdot 10^{12}$ Hz. Über die Wahl von a und $\nu(\lambda)$ siehe die folgenden Paragraphen.

Der parabelähnliche Bogen in Abb. 3 gibt die Spektralfarben und die seine Endpunkte verbindende Strecke die „Purpurgerade“. Gegenüber diesem Spektralfarbenkreise und dem Ostwaldschen Farbenkreise (punktierte Kurve) zeichnet sich der Normalfarbenkreis durch die denkbar einfachste geschlossene Kurve aus. Man beachte dabei, daß eine Ellipse wegen der jederzeit noch möglichen linearen Koordinatentransformation auf Kreisform gebracht werden kann (m. a. W.: wegen der Projektivität der gesamten Darstellung ist die Ellipse dem Kreise gleichwertig). In diesem Sinne ist ein Normalfarbenkreis also wirklich ein Farbenkreis im Unterschied zu den anderen „Farbenkreisen“!

Wichtig ist, daß diese Bevorzugung der Normalfarbenkreise nicht an die König-Dieterichschen Grundempfindungskurven, ja nicht einmal an die Dimension 3 gebunden ist, wie oben bereits bei der allgemeineren Betrachtung ausinandergesetzt wurde. Vom Standpunkte einer physikalischen Farbenlehre haben gerade solche Zusammenhänge besondere Bedeutung, die kovariant gegen beliebige Koordinatentransformation und unabhängig von der gewählten Dimension sind.

Wenn die geläufige Farbenlehre die Dimensionszahl 3 wählte, so geschah das aus einem physiologischen Grunde. Äquivalenzklassen zur Dimension 3 entsprechen in ihrer Mannigfaltigkeit gerade dem Unterscheidungsvermögen des menschlichen Auges bei Hellanpassung; das Unterscheidungsvermögen bei Dämmerungssehen entspricht der Dimension 1.

3 ist ferner die kleinste Dimension, für die eine Äquivalenztheorie nicht zum Trivialien entartet, und zugleich die größte Dimension, für die der Raum R unserer Anschauung unmittelbar zugänglich ist. Das gibt der dreidimensionalen Farbenlehre eine Vorzugsstellung auch ohne physiologische Gesichtspunkte.

Wollten wir alle nach Df. 1 definierten Farben unterscheiden können (dies soll heißen, daß verschiedene Farben nicht äquivalent sein sollten), so würde als Menge \mathfrak{M} nur ein Kontinuum ausreichen; wir müßten dann dem Σ -Zeichen in der oben gegebenen Darstellung zweckmäßig die Bedeutung eines Integrals zuweisen. Beschränken wir uns auf analytische Funktionen als Farben, so können wir mit abzählbar vielen Dimensionen auskommen. Immer freilich haben wir bei unendlich vielen Dimensionen Vorsichtsmaßregeln und Einschränkungen nötig; so ist es z. B. zweckmäßig, als Grundfunktion $\nu(\lambda)$ nur eine beschränkte Funktion zuzulassen. (Siehe § 8!) Um die gebräuchliche Farbenlehre als Teil einer physikalischen Farbenlehre im Raum mit abzählbar vielen Dimensionen zu erhalten, wählt man zweckmäßig eine Basis $b_i(\nu)$, $i = 1, 2, 3, \dots$ mit $b_1(\nu(\lambda)) = \bar{y}(\lambda)$; $b_2(\nu(\lambda)) = \bar{z}(\lambda)$; $b_3(\nu(\lambda)) = \bar{x}(\lambda)$; darin sind \bar{x} , \bar{y} , \bar{z} die IBK-Normalreizbeträge des Normblatts DIN 5053. Als b_4 sei die Augenempfindlichkeitsfunktion des Dämmerungssehens vorgeschlagen; für physiologische Fragen hat dann hauptsächlich die Projektion aller Farben auf diesen vierdimensionalen Unterraum Bedeutung.

Unter den dreidimensionalen Farbenlehren spielt für die physikalische Meßtechnik noch die auf dem Agfa-Color-Verfahren basierende eine gewisse Rolle;

wählen wir die drei Grundempfindlichkeitskurven des Agfa-Color-Films als Basis, so schreiben wir die Äquivalenz

$$f_1(v) \stackrel{AC}{43964} f_2(v)$$

17

unter Angabe der Emulsionsnummer und einer Kennnummer für die Filmentwicklung. Bezeichnet man Farben als äquivalent, die dem Auge ununterscheidbar sind, so notiert man zweckmäßig unter dem Äquivalenzzeichen den Namen des Beobachters; leider genügt auch das nicht immer, da das Urteil „ununterscheidbar“ auch von der „Vorgeschichte“ des Auges abhängt⁷⁾.

Bei Verwendung der Normalfarben in der Meßtechnik kann auf Äquivalenzbetrachtungen völlig verzichtet werden; gerade das war die Absicht bei Einführung der Normalfarben. Die Betrachtungen dieses Paragraphen sollten vor allem die Einordnung der Normalfarben in eine verallgemeinerte, aber der üblichen analoge Farbenlehre zielen.

§ 7. Experimentelle Realisierungen von Normalfarbenkreisen

Eine experimentelle Realisierung von Normalfarben mit einer Schablone im Spektrum entsprechend der Abb. 1 hat sich besonders bewährt, wenn es auf mehr prinzipielle Untersuchungen über die physiologische Wirkung der zu verschiedenen $a \cdot v(\lambda)$ gehörigen Normalfarbensysteme ankam.

Auf eine elegantere Realisierung für meßtechnische Zwecke deutet der Cosinus in der die Normalfarben charakterisierenden Funktion (10). Die an einer Planparalleplatte der Dicke d senkrecht reflektierte Intensität ist

$$I_0 \cdot \cos^2 \left(2\pi \frac{d}{c} \nu + \varphi \right); \quad (24)$$

darin ist I_0 die einfallende Intensität, ν die Frequenz, c die Lichtgeschwindigkeit und φ eine Phasenverschiebung. Die beiden miteinander interferierenden Wellen sind als gleich stark vorausgesetzt. Die Farben dünner Blättchen sind also Normalfarben; doch gehören die bei verschiedener Dicke d erscheinenden Farben zu verschiedenen Normalfarbensystemen. Man erhält den Farbenkreis eines einheitlichen Systems ersichtlich (vergleiche (10)!) durch Variation der Phasenverschiebung von 0 bis 2π unter Konstanthalten der Dicke d . Das freilich erwies sich in der Praxis als schwierig; es erfordert die Bedampfung der Unterlage und der Oberfläche einer dünnen Zaponlackhaut (auf Glas oder dergleichen) mit mindestens drei verschiedenen Metallen in veränderlicher Dicke. Dennoch ist die Frequenz durch ihr Auftreten in der eine Welle darstellenden Funktion so ausgezeichnet, daß man ihr vor vielen anderen möglichen Grundfunktionen, jedenfalls vor λ selbst, den Vorzug geben wird.

Aber selbst dann, wenn wir über $\nu(\lambda)$ verfügen, indem wir es mit der Frequenz identifizieren, bleibt das a , d. h. die Periodenlänge $g = \frac{1}{a}$ noch offen. Dem entspricht bei der Realisierung mittels Farben dünner Blättchen die Wählbarkeit der

⁷⁾ Siehe § 1.

Dicke $d = \frac{1}{2} a \cdot c$. Unter den hiernach noch möglichen Systemen zeichnet sich jedoch eines aus, das besonders günstig zur Länge des sichtbaren Spektralbereichs liegt. Während sehr große d (d. h. kleine Periode g) bekanntlich nur wenig von Grau abweichend scheinende Farben (sogenanntes Weiß höherer Ordnung) liefern, muß offenbar ein allzu kleines d dazu führen, daß in Gl. (24) die \cos^2 -Funktion für alle Frequenzen des Sichtbaren einen etwa einheitlichen Wert hat, also praktisch etwa Graustufen zwischen Weiß und Schwarz durchläuft, wenn μ geändert wird; das erhellt aus einem Verschieben einer \cos^2 -Schablone mit sehr großem g (im Vergleich zur Breite des sichtbaren Spektrums) ebenso, wie bei kleinem g das Weiß höherer Ordnung aus der Schablonenfigur anschaulich wird.

Zwischen diesen Extremen liegt ein Optimum bei $d = 0,5 \cdot 10^{-4}$ cm (d. h. $g = 300 \cdot 10^{12}$ Hz). Der Farbenkreis der gesättigten Normalfarben dieses Systems umfaßt dann die Farben Rot, Gelb, Grün, Blau, Violett und Purpur in einer nicht nur kontinuierlichen sondern auch dem Auge überall hinreichend gesättigt erscheinenden Form. Es handelt sich hier natürlich um eine Entscheidung von der Physiologie her, die noch im Paragraphen 8 näher beleuchtet werden soll. Doch beachte man, daß hier von der Physiologie keine Äquivalenzbetrachtung, sondern ausschließlich eine Angabe über die Länge des sichtbaren Spektrums stammt.

Eine dritte Gruppe von Realisierungen der Normalfarben schließt an die bekannten Farben an, die man an optisch anisotropen oder optisch aktiven Stoffen zwischen Polarisatoren beobachtet. Alle diese Farben sind Normalfarben; das Problem liegt für unsere Aufgabe meist nur darin, Vorrichtungen zu finden, die Normalfarben eines einheitlichen Systems (d. h. einer einheitlichen Funktion $a \cdot v(\lambda)$ in Gl. (10)) liefern.

Am leichtesten gelingt das unter Verwendung der Rotationsdispersion mit einem Verfahren, das schon von Helmholtz im Leukoskop verwendet, und dessen spezielle technische Zurichtung für die Zwecke des Farbschlierenverfahrens in der oben genannten Arbeit^{*)} beschrieben wurde. Hier soll es vor allem im Hinblick auf die mit ihm realisierbaren Grundfunktionen $a \cdot v(\lambda)$ untersucht werden.

Eine d cm dicke Platte optisch aktiven Materials mit einem spezifischen Drehvermögen $[\alpha(\lambda)]$ liege zwischen einem festen Polarisator und einem drehbaren Analysator, dessen Polarisationsrichtung mit der des Polarisators den veränderlichen Winkel β bildet. Der von der gesamten Vorrichtung hindurchgelassene Bruchteil der einfallenden Intensität (natürlichen Lichtes) ist dann

$$\cos^2(\pi d [\alpha(\lambda)] - \beta). \quad (25)$$

Für jede Stellung des Analysators erhalten wir also eine Normalfarbe. Alle diese Normalfarben gehören zu einem einheitlichen System, das durch $a \cdot v(\lambda) = d \cdot [\alpha(\lambda)]$ charakterisiert ist. Wählen wir also irgendein Material, so bleibt nur noch über die Dicke d zu entscheiden. Das kann von der Physiologie her wie bei den Farben dünner Blättchen geschehen und führt z. B. bei Quarz (Kristallachse in Richtung der optischen Achse) zu einer 4 mm dicken Platte. In der oben genannten Arbeit^{*)} war in Abb. 6 der Verlauf von $[\alpha(\lambda)]$ in Abhängigkeit von der Frequenz angegeben worden. Da die Kurve nur wenig von einer Geraden abweicht, ist das durch

^{*)} Ann. Physik 1. c.

Quarz⁹⁾ festgelegte Normalfarbensystem nicht sehr verschieden von dem Frequenzsystem, auf das die Farben dünner Blättchen geführt hatten. Auch ein reines Frequenzsystem läßt sich mit Hilfe der Rotationsdispersion aufbauen, wenn man die Krümmung in der zu $[\alpha(\lambda)]$ gehörigen Kurve durch Kombination mit einem geeigneten anderen optisch aktiven Stoff kompensiert.

Wegen der großen Auswahl optisch aktiver Stoffe kann man aber auch praktisch alle gewünschten Normalfarbensysteme mit fast beliebig gewähltem $a \cdot \nu(\lambda)$ realisieren. Wir werden im nächsten Paragraphen dieser Arbeit noch Gebrauch davon machen.

§ 8. Zur Entscheidung über die Grundvariable

Aus unseren Forderungen der Periodizität und Additivität war nur zu folgern, daß die Normalfarben die Form

$$F(\lambda) = A + B \cos^2(\pi a \{ \nu(\lambda) - \nu(\lambda_0) \}); \quad \mu = \nu(\lambda_0)$$

haben müssen. Die Grundvariable $\nu(\lambda)$ und die Zahl a waren dabei noch verfügbar; von $\nu(\lambda)$ war lediglich verlangt worden, daß es eine stetige, eigentlich monotone Funktion von λ ist. So haben wir es also nicht mit einem Normalfarbensystem, sondern mit einer Menge von Normalfarbensystemen zu tun, deren Kardinalzahl gleich der Mächtigkeit der Menge aller stetigen Funktionen ist.

Solange noch Lösungen unseres Problems in solcher Vielfalt gleichberechtigt nebeneinander stehen, könnte man die Aufgabe zur Schaffung eines rationellen Farbenkreises als nicht restlos gelöst betrachten.

Man kann aber leicht einsehen, daß die verbliebene Freiheit nur Vorteile bietet und uns erlaubt, jeweils ein Normalfarbensystem auszusuchen, wie es gerade einer speziellen praktischen Anwendung besonders angemessen ist. Neben den mit Farben dünner Blättchen besonders leicht realisierbaren Frequenzsystemen ($\nu(\lambda) = \frac{c}{\lambda}$) und den Quarzsystemen usw. wollen wir noch drei Systeme näher betrachten.

a) Die Temperatursysteme

Wenn wir als Beleuchtung die Strahlung des schwarzen Körpers von der Temperatur T benutzen, und wenn wir auf sie eine Schablonenfunktion $\cos^2(\pi a(\nu - \mu))$ mit ν als Frequenz anwenden, so werden offenbar die so entstehenden gesättigten reduzierten Normalfarben eine von μ abhängige resultierende Gesamtintensität liefern. Verlangen wir dagegen über die Additivität hinaus noch eine für alle reduzierten Normalfarben des Systems konstante Gesamtintensität, so ist das nur zu realisieren bei Transformation auf eine Grundvariable $\nu(\lambda)$, über der als Abszisse nun die Kirchhoff-Planck-Funktion $E(\lambda; T)$ eine Konstante wird. Da die Intensität im Spektrum von λ bis $\lambda + d\lambda$ gleich $E(\lambda; T)$ ist, so muß also die neue Grundvariable die Differentialgleichung

$$d\nu(\lambda) = K_1 \cdot E(\lambda; T) d\lambda \quad (26)$$

⁹⁾ Daß den „Quarzfarben“ bei Drehung des Analysators eine Ellipse im Farbedreieck entspricht, entdeckte S. Rösch unabhängig vom Verfasser und gleichzeitig.

erfüllen; Integration ergibt

$$\nu(\lambda) = K_1 \cdot \int_0^\lambda E(\lambda'; T) d\lambda' + K_2. \quad (27)$$

Da $E(\lambda; T) > 0$ für alle λ ist, so ist $\nu(\lambda)$ eine stetige eigentlich monotone Funktion von λ , wie in Df. 1 verlangt worden war. ν geht nur in Differenzform in die Formeln ein, wir können daher o. B. d. A. $K_2 = 0$ setzen. Wir wählen ferner $K_1 = 1$ o. B. d. A., da a ohnehin noch verfügbar bleibt. Die Gesamtintensität des von der reduzierten Normalfarbe (13) ausgehenden Lichtes bei der vorgegebenen Beleuchtung ist

$$\begin{aligned} I(\lambda_0) &= \frac{1-B}{2} + B \cdot \int_0^\infty E(\lambda; T) \cos^2(\pi a \cdot \{\nu(\lambda) - \nu(\lambda_0)\}) d\lambda \\ &= \frac{1-B}{2} + B \cdot \int_0^{v(\infty)} \cos^2(\pi a \{v - \nu(\lambda_0)\}) dv \end{aligned} \quad (28)$$

und kann dann in der Tat von λ_0 unabhängig sein, sofern nur $v(\infty)$ ein ganzes Vielfaches von $g = \frac{1}{a}$, also $v(\infty) = n g$ ist; dabei ist

$$v(\infty) = \int_0^\infty E(\lambda; T) d\lambda. \quad (29)$$

Wir haben also die geforderte Konstanz der Gesamtintensität über alle reduzierten Normalfarben gesichert, wenn wir über g und v in dem angegebenen Sinne verfügen. Die reduzierten Normalfarben „zur Temperatur T “ nehmen damit die Form an

$$F_T(\lambda) = \frac{1-B}{2} + B \cos^2 \left(\pi n \frac{\int_{\lambda_0}^\lambda E(\lambda'; T) d\lambda'}{\int_0^\infty E(\lambda'; T) d\lambda'} \right); \quad (30)$$

n ist darin eine beliebige natürliche Zahl. Einen Gesichtspunkt, der für $n = 1$ spricht, diskutieren wir in § 9.

b) Systeme zu vorgegebenem Beleuchtungsspektrum

Wählt man statt der Strahlung des schwarzen Körpers z. B. das Sonnenspektrum als Beleuchtung, so findet man in völlig entsprechender Weise das Normalfarbensystem, das bei Sonnenbeleuchtung intensitätsgleiche Farben enthält. Ähnliches gilt für jedes Spektrum; in (30) ist $E(\lambda; T)$ dann nur als das jeweils gewählte Beleuchtungsspektrum zu deuten.

c) Das Normalfarbensystem gleicher physiologischer Helligkeit

Verlangen wir, daß alle reduzierten Normalfarben bei Beleuchtung mit dem Spektrum $S(\lambda)$ dem Auge gleich hell erscheinen, so müssen wir die Augenempfind-

lichkeitskurve $V(\lambda)$ berücksichtigen und statt Gl. (30) schreiben

$$F_V(\lambda) = \frac{1-B}{2} + B \cos^2 \left(\pi \frac{\int_{\lambda_0}^{\lambda} V(\lambda') S(\lambda') d\lambda'}{\int_0^{\lambda_0} V(\lambda') S(\lambda') d\lambda'} \right). \quad (31)$$

Über die in Gl. (30) noch auftretende natürliche Zahl n ist in diesem physiologischen Normalfarbensystem bereits zwanglos durch die Festsetzung $n=1$ verfügt worden. Mit Hilfe der Schablonenvorrichtung überzeugt man sich leicht, daß größere n zu Farben führen, die sich von einem Weiß höherer Ordnung nicht mehr wesentlich unterscheiden. Für $n=1$ liefert das System gesättigte und in jeder Weise befriedigende gleich helle Normalfarben. (Siehe auch § 9!)

Das System (31) kann durch Kombination optisch aktiver Stoffe zwischen Polarisatoren mit großer Annäherung realisiert werden. Da es auch für die Farbenlehre Bedeutung hat, wird es in einer gesonderten Veröffentlichung näher betrachtet.

Wenn hier ohne Zweifel ein physiologischer Gesichtspunkt in unsere Betrachtungen eingebrochen ist, so geschah das erstens nur im Zusammenhang mit einem noch verfügbaren Freiheitsgrad und daher ohne Verstoß gegen die ursprünglichen physikalischen Forderungen. Zweitens bedeutet die Augenempfindlichkeitskurve lediglich eine präzisiertere Angabe über den sichtbaren Spektralbereich, der noch weniger Willkür anhaftet als der Festlegung durch zwei Sichtbarkeitsgrenzen.

Für die physikalischen Anwendungen werden wir neben diesem System vor allem das Quarzsystem und das Frequenzsystem benutzen.

§ 9. Darstellung jeder vorgegebenen Farbe mittels Normalfarben

Multiplikative Mischung gesättigter reduzierter Normalfarben mit sich selbst ergibt höhere Potenzen einer \cos^2 -Funktion. Der an $\nu = \mu$ vorhandene Maximalwert 1 bleibt dabei erhalten, während der Funktionswert an allen anderen Stellen gegen Null geht, wenn die multiplikative Mischung fortgesetzt wird. Durch kombinierte multiplikative und additive Mischung kann man daher aus Normalfarben (10) jede beliebige Farbe mit der Periode g beliebig gut approximieren.

Ist die Grundfunktion $\nu(\lambda)$ für das gesamte Intervall $0 \leq \lambda < \infty$ beschränkt, so haben wir bezüglich der Variablen ν nur ein endliches Intervall $\nu(0) \leq \nu < \nu(\infty)$ vor uns. Die Normalfarben

$$A + B \cos^2 \left(\pi \frac{\nu(\lambda) - \nu(\lambda_0)}{\nu(\infty) - \nu(0)} \right), \quad A \geq 0; B \geq 0; A + B \leq 1, \quad (32)$$

die dem System (30) analog gebildet sind, nehmen den Wert 1 höchstens für ein λ an. Das Normalfarbensystem zu einer beschränkten Grundfunktion läßt daher jede vorgelegte Farbe mit jeder gewünschten Genauigkeit durch kombinierte multiplikative und additive Mischung dieser Normalfarben approximieren. Auch das ist ein wichtiger Gesichtspunkt zur Wahl von $n=1$ im letzten Paragraphen, ohne daß hierbei physiologische Fragen eine Rolle spielen.

Diese Darstellbarkeit aller Farben trifft insbesondere für das Normalfarbensystem gleicher Helligkeit (31) zu, nicht aber z. B. für das Frequenzsystem, dessen Grundfunktion nicht beschränkt ist.

§ 10. Messung von Normalfarben

Das innerhalb eines jeden Normalfarbensystems für eine Normalfarbe charakteristische Zahlentripel A, B, λ_0 läßt sich verhältnismäßig einfach messen mit einem etwas vervollständigten Taschenspektroskop. In ihm seien übereinander das Spektrum des von der Farbe beeinflussten und des direkten Lichtes beobachtbar. Das direkte Licht wird durch einen vor dem Spalt verschieblichen geeichten Graukeil meßbar geschwächt, bis das Maximum des Normalfarbenspektrums gleiche Helligkeit wie die entsprechende Stelle des Vergleichsspektrums zeigt; aus der Graukeilstellung entnimmt man den Wert $A + B$. Ebenso liefert eine Einstellung des Vergleichsspektrums auf gleiche Helligkeit mit dem Minimum den Wert von A . Der Ort im Spektrum, an dem das Maximum liegt, ergibt das λ_0 . Mit einem handelsüblichen Taschenspektroskop von I. D. Moeller konnten λ_0 auf einige $m\mu$ und A, B auf einige Prozent genau gemessen werden.

Zusammenfassung

Die in § 3 eingeführten Normalfarben zeichnen sich durch folgende Eigenschaften aus:

1. Gesättigte Normalfarben liefern einen Farbenkreis, der dem Ostwaldschen ähnelt; doch bietet ein Normalfarbenkreis außer der Kontinuierlichkeit den Vorteil, daß additive Mischung von Normalfarben stets wieder eine Normalfarbe im Sinne der strengen physikalischen Gleichheit ergibt. Daher werden subjektive Urteile („physiologisch äquivalent“) überflüssig, wenn ein Normalfarbenkreis statt des Ostwaldschen (oder anderer) in der physikalischen Meßtechnik benutzt wird.

2. In einer allgemeineren Äquivalenztheorie der Farben wird ein Normalfarbenkreis stets durch eine Ellipse dargestellt, während dem Ostwaldschen Farbenkreise eine unübersichtliche, komplizierte und willkürlich erscheinende Kurve entspricht.

3. Die Normalfarben sind einfach mit physikalischen Hilfsmitteln zu realisieren und genau reproduzierbar.

4. Das eine Normalfarbe charakterisierende Zahlentripel kann einfach und objektiv (ohne Eingehen von physiologischen Momenten) gemessen werden.

5. Es gibt unendlich viele Normalfarbensysteme; und zwar bleibt eine eigentlich monotone stetige Funktion („Grundfunktion“) verfügbar. Diese kann der jeweils vorliegenden Aufgabe optimal angepaßt werden. So kann man z. B. zu einer gegebenen Beleuchtung eine Grundfunktion so wählen, daß nach Passieren der die Normalfarben realisierenden Vorrichtung gleiche Gesamtintensität (oder nach Wunsch auch gleiche visuelle Helligkeit) über den ganzen Farbenkreis resultiert. Die Eigenschaften 1 bis 4 bleiben auf jeden Fall von dieser Wahl unberührt.

6. Bei einem Normalfarbensystem zu beschränkter Grundfunktion läßt sich jede beliebig gegebene Farbe durch kombinierte multiplikative und additive Mischung der Normalfarben beliebig genau in ihrem spektralen Verlauf approximieren.

Dem Direktor des Instituts, Herrn Professor Dr. Lochte-Holtgreven, dankt Verfasser für freundliche Unterstützung und Herrn stud. phil. Steinebach für wertvolle Hilfe bei numerischen Rechnungen.

Kiel, Institut für Experimentalphysik der Universität.

(Bei der Redaktion eingegangen am 27. Mai 1950.)

Untersuchungen über die Beeinflussung der spektralen Empfindlichkeit von Selen-Photoelementen durch Photostrom und Außenwiderstand

Von Irmgard Wolf

(Mit 7 Abbildungen)

Inhaltsübersicht

I. Für die Messung der spektralen Empfindlichkeit von Selen-Photoelementen werden zwei Methoden beschrieben:

1. Aufnahme der spektralen Empfindlichkeit unter Verwendung eines äquiennergetischen Spektrums,

2. Aufnahme der spektralen Empfindlichkeit bei Konstanthaltung des Photostroms.

II. Es wird die Abhängigkeit der spektralen Empfindlichkeit im 1. Fall von der Größe der eingestrahnten Intensität, im 2. Fall von der Größe des Photostroms, in beiden Fällen von dem angeschlossenen Außenwiderstand, nachgewiesen.

III. Als Ursache dafür wird das für die einzelnen Wellenlängen verschiedene nichtproportionale Anwachsen des Photostroms mit der eingestrahnten Lichtintensität erkannt.

Barnard¹⁾ vertritt in einer Arbeit über die spektrale Empfindlichkeit von Selen-Photoelementen die Ansicht, daß diese von den Betriebsbedingungen, insbesondere von dem durch das Element fließenden Photostrom und dem Außenwiderstand abhängig ist.

Wir wollen die spektrale Empfindlichkeitsverteilung eines Photoelements im Anschluß an die übliche Meßmethode definieren als die Abhängigkeit des Photostroms von der Wellenlänge des eingestrahnten Lichtes, unter der Voraussetzung, daß dieses ein äquiennergetisches Spektrum hat. Die relative spektrale Empfindlichkeitsverteilung φ_λ erhält man, wenn man das Maximum der eben erwähnten Kurve 1 setzt und die übrigen Werte entsprechend reduziert.

Ist nun die Form von φ_λ durch eine Messung bei einer beliebigen Energie des eingestrahnten äquiennergetischen Spektrums eindeutig bestimmt, oder ergeben sich für verschiedene Energiewerte verschiedene Kurvenformen? Das letztere würde also bedeuten, daß φ_λ von der Energie des eingestrahnten Lichtes und somit auch von dem durch das Element fließenden Photostrom abhängig ist, denn der Photostrom ist ja eine Funktion der eingestrahnten Lichtintensität, bzw. der Beleuchtungsstärke des auffallenden Lichtes, wenn das Element voll ausgeleuchtet ist,

¹⁾ G. P. Barnard, Proc. phys. Soc. 51, 284 (1939).

wie es bei den hier vorliegenden Messungen stets der Fall war. Eine Antwort auf diese Frage wäre prinzipiell möglich durch Aufnahme der Kurven φ_λ bei verschiedener Energie. Dies ist jedoch meßtechnisch schwer zu verwirklichen, da die Herstellung monochromatischen Lichtes großer Intensität mit Schwierigkeiten verbunden ist. Es werden aber große Intensitätsunterschiede benötigt, um eine eventuelle Änderung von φ_λ feststellen zu können, denn bei bloßer Verdopplung der Intensität z. B. kann eine Kurvenänderung im Rahmen der Meßgenauigkeit nicht festgestellt werden.

Gleiche Kurvenform von φ_λ für verschiedene Energien bedeutet je eine proportionale Änderung des Photostroms mit der eingestrahnten Intensität, für jede Wellenlänge mit dem gleichen Proportionalitätsfaktor. Diese Proportionalität bei kleinen Energieänderungen wird bei der Aufnahme der Kurve φ_λ benutzt, da man ein äquiengetisches Spektrum für einen großen Wellenlängenbereich nicht herstellen kann.

Wenn aber z. B. bei einer Verzehnfachung der eingestrahnten Intensität etwa im roten Licht der Photostrom sich nicht mehr verzehnfacht, sondern nur noch das 9,5fache des ursprünglichen Wertes beträgt, im blauen dagegen nur noch etwa das 9fache, so entsteht durch diese Abweichung von der Proportionalität natürlich eine völlig veränderte Form der Kurve φ_λ und somit eine Abhängigkeit der Form von der eingestrahnten Intensität.

In dieser Weise läßt sich aber der Versuch nicht für das ganze Spektrum durchführen, so daß man Aussagen über eine etwaige Änderung von φ_λ nur auf indirektem Wege über die Messung des Durchlaßkoeffizienten von farbigen Gläsern erhalten kann. Dieser Weg ist von Barnard¹⁾ und Dresler²⁾ beschritten worden. Eine in der Arbeit von Dresler vorhandene Unstimmigkeit gab den Anlaß, in der vorliegenden Arbeit das Problem nochmals zu untersuchen.

Der Durchlaßkoeffizient eines farbigen Glases gibt an, welcher Bruchteil der ankommenden Lichtintensität insgesamt durch das Filter hindurchgelassen wird. Dieser Bruchteil ist direkt durch das Verhältnis der Beleuchtungsstärken — in einer festen Entfernung von der Lichtquelle — im gefilterten und ungefilterten Licht gegeben. Die Beleuchtungsstärke gibt ja die pro cm und sek. auffallende Energie der Strahlung an. Hat man einen Strahlungsempfänger, der direkt diese Beleuchtungsstärken zu messen gestattet, so kann man also den Durchlaßkoeffizienten durch das Verhältnis der Beleuchtungsstärken in fester Entfernung ermitteln. Bei einem Strahlungsempfänger, der nur die Gleichheit zweier Beleuchtungsstärken registrieren kann (wie z. B. das menschliche Auge), ergibt sich — wie man leicht einsieht — der Durchlaßkoeffizient zu r_2^2/r_1^2 (r_1 Entfernung des Strahlungsempfängers von der Lichtquelle ohne, r_2 mit eingeschaltetem Farbfilter: in beiden Fällen gleiche Beleuchtungsstärke).

Es ist klar, daß der so bestimmte Durchlaßkoeffizient außer von der Energieverteilung der benutzten Lichtquelle (d. h. von der Temperatur im wesentlichen) auch von den selektiven Eigenschaften des benutzten Strahlungsempfängers abhängt. Für das Auge als Strahlungsempfänger ergibt sich dann der Durchlaßkoeffizient eines farbigen Glases als

$$D = \frac{\int_0^\infty \tau_\lambda E_\lambda V_\lambda d\lambda}{\int_0^\infty E_\lambda V_\lambda d\lambda},$$

²⁾ A. Dresler, „Das Licht“ 11, H. 7 (1941).

wobei τ_λ die spektrale Durchlässigkeit des Glases bezogen auf äquiennergetisches Spektrum, E_λ die relative Energieverteilung der Lichtquelle und V_λ die relative spektrale Hellempfindlichkeit des Auges, jeweils für die Wellenlänge λ bedeuten.

Bestimmen wir den Durchlaßkoeffizienten mit Hilfe eines Photoelementes, so wird entsprechend

$$D^* = \frac{\int_0^\infty \tau_\lambda E_\lambda \varphi_\lambda d\lambda}{\int_0^\infty E_\lambda \varphi_\lambda d\lambda}, \quad (1)$$

wobei φ_λ die relative spektrale Empfindlichkeit des Photoelementes bezeichnet.

Die Integrale in (1) bedeuten offenbar die Gesamtwirkung der eingestrahnten Lichtintensität auf das Photoelement (im Zähler mit, im Nenner ohne eingeschaltetes Farbfilter), oder mit anderen Worten die von dem Photoelement bewerteten Beleuchtungsstärken des gefilterten bzw. ungefilterten Lichtes in fester Entfernung von der Lichtquelle. Unter der Voraussetzung, daß der Photostrom für jede Wellenlänge der eingestrahnten Lichtenergie proportional ist und sich die Wirkungen der verschiedenen Wellenlängen addieren, ist der gesamte Zellenstrom den Integralen in Formel (1) proportional. Der Durchlaßkoeffizient läßt sich also dann bestimmen als der Quotient der Photoströme bei Bestrahlung des Photoelementes mit gefiltertem und ungefiltertem Glühlampenlicht, wobei die Entfernung von der Lichtquelle konstant gehalten wird.

Nimmt man an, daß die Kurve φ_λ für ein Photoelement nur eine einzige Form hat, unabhängig von der jeweiligen Energie des bei der Aufnahme benutzten äquiennergetischen Spektrums, so erkennt man aus Formel (1), daß D^* für ein bestimmtes Element durch die Energieverteilung der Lampe (also ihre Temperatur) und das jeweils benutzte Filter gegeben ist. Insbesondere ist D^* also unabhängig von der Entfernung des Photoelementes von der Lichtquelle.

Tatsächlich ergibt die Messung von D^* verschiedene Werte des Durchlaßkoeffizienten für verschiedene Entfernungen. Was ist die Ursache für diese Erscheinung? Es liegt nahe, hierfür das bereits oben als möglich angenommene nicht-proportionale Anwachsen des Photostroms mit der eingestrahnten Intensität verantwortlich zu machen. Für weißes Licht geht aus den bekannten „Charakteristiken“ von Selen-Photoelementen (Abhängigkeit des Photostroms von der eingestrahnten Intensität bzw. der Beleuchtungsstärke) hervor, daß diese Kurven für größere Beleuchtungsstärken immer mehr vom geraden Verlauf abweichen. Dies Verhalten tritt um so stärker in Erscheinung, je größer der angeschlossene Außenwiderstand ist. Für monochromatisches Licht verhält sich das Photoelement genau so, und zwar ergibt sich für jede Wellenlänge ein anderes nicht-proportionales Verhalten (wie aus den Messungen, die weiter unten besprochen werden sollen, geschlossen werden kann). Daß derartige Eigenschaften der Photoelemente eine Gestaltsänderung der φ_λ -Kurve bedingen, hatten wir bereits oben gesehen.

Man könnte nun versuchen, diese verschiedenen Formen der φ_λ -Kurven aus den verschiedenen, nach der 1. Methode gemessenen, Durchlaßkoeffizienten nach Formel (1) zu berechnen, d. h. für jeden gemessenen Wert von D^* eine φ_λ -Kurve zu konstruieren derart, daß sich mit ihr aus (1) gerade der gemessene Wert D^* berechnen läßt. Ein solches Vorgehen wäre aber begrifflich nicht ganz einwandfrei.

Man kann sich leicht überlegen, daß man auf diese Weise keineswegs etwa die oben erwähnten φ_λ -Kurven für verschiedene Energien des eingestrahnten äqui-

energetischen Spektrums erhält, sondern gewissermaßen eine „Arbeitskurve“ der spektralen Empfindlichkeit, die durch die absolute Energieverteilung \bar{E}_λ der benutzten Beleuchtungsquelle bestimmt wird, ja man sieht sogar sofort, daß in den Zähler und Nenner von (1) verschiedene φ_λ einzusetzen sind, denn für ungefiltertes Licht ist die auf das Element wirkende Energie ihrer Größe nach durch \bar{E}_λ gegeben, während für Bestrahlung mit gefiltertem Licht die tatsächlich auf das Element auffallende Energie nur $\tau_\lambda \bar{E}_\lambda$ beträgt. Eine Berechnung von φ_λ aus D^* nach (1) ist also nicht möglich.

(Wenn die φ_λ -Kurven für verschiedene Energien eines äquinergetischen Spektrums vorlägen, könnte man sich die „Arbeitskurve“ des Photoelementes

für die vorliegende Energieverteilung \bar{E}_λ bzw. $\tau_\lambda \bar{E}_\lambda$ konstruieren. Nach Abb. 1 wählt man dazu aus der Schar der φ_λ -Kurven für die Wellenlänge λ die zu dem Energiewert \bar{E}_λ (bzw. $\tau_\lambda \bar{E}_\lambda$) gehörige Kurve aus. Ihre Ordinate an der Stelle λ gibt den Ordinatenwert der nicht-reduzierten „Arbeitskurve“. Man sieht, daß dies Verfahren eine gewisse Ähnlichkeit mit der Konstruktion der Arbeitskennlinien einer Verstärkerröhre aus den statischen Kennlinien hat. Die Bezeichnung „Arbeitskurve“ erscheint daher naheliegend.)

Wir haben dann aber noch die für jeden Strahlungsempfänger vorhandene

2. Möglichkeit zur Bestimmung des Durchlaßkoeffizienten, indem wir bei der Messung die Beleuchtungsstärke konstant halten und D^* als das umgekehrte Verhältnis der Entfernungsquadrate ermitteln. Es ist naheliegend, daß man bei dem Photoelement gleiche Beleuchtungsstärke durch gleichen Photostrom bestimmt.

Wir halten also bei dieser Messung den Photostrom konstant, und es ergibt sich nach dem oben Gesagten

$$D^* = \frac{r_2^2}{r_1^2} \quad (2)$$

Wenn der Photostrom der eingestrahnten Lichtintensität für alle Wellenlängen streng proportional ist, so ist der nach beiden Methoden gemessene Durchlaßkoeffizient der gleiche.

Wollen wir nun den nach der 2. Methode gemessenen Durchlaßkoeffizienten aus Formel (1) berechnen, so müssen wir die Definition der relativen spektralen Empfindlichkeit abändern. Hatten wir oben φ_λ definiert als Abhängigkeit des Photostroms von der Wellenlänge eines eingestrahnten äquinergetischen Spek-

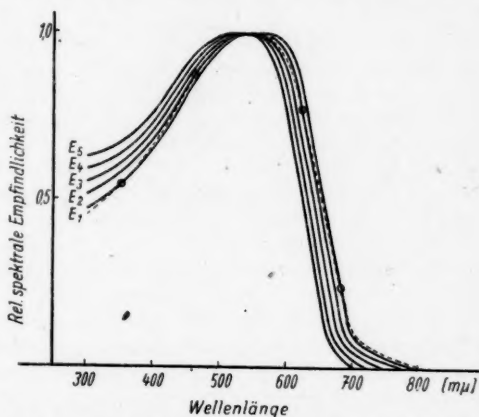


Abb. 1. Konstruktion der „Arbeitskurve“ der spektralen Empfindlichkeit aus den für konstante Energien $E_1 > E_2 > E_3 > E_4 > E_5$ aufgenommenen φ_λ -Kurven

trums, so müssen wir jetzt dafür Sorge tragen, daß bei der Aufnahme der φ_λ -Kurve der Photostrom konstant bleibt. Wir werden also die Messung der φ_λ -Kurve folgendermaßen durchführen: Wir bestrahlen das Photoelement mit dem Licht einer beliebigen Wellenlänge λ_0 und erzeugen durch eine Veränderung der Intensität des eingestrahnten Lichtes einen bestimmten Photostrom i . Die auf das Element auffallende gesamte Lichtintensität der Wellenlänge λ_0 betrage E_0 (man kann sie z. B. durch ein Thermoelement messen). Für jede andere Wellenlänge λ wird dann wiederum die Intensität des eingestrahnten Lichtes so eingestellt, daß der erzeugte Photostrom gerade wieder i beträgt. Ist die hierfür benötigte Lichtintensität E , so gilt

$$\frac{y}{y_0} = \frac{E_0}{E},$$

wobei y bzw. y_0 die Ordinaten der spektralen Empfindlichkeitskurve für die Wellenlängen λ bzw. λ_0 sind. In der so erhaltenen Kurve der spektralen Empfindlichkeit setzt man das Maximum 1 und rechnet die anderen Werte entsprechend um und bekommt dann die Kurve der relativen spektralen Empfindlichkeitsverteilung φ_λ für den konstanten Photostrom i . Besteht Proportionalität zwischen eingestrahelter Lichtintensität und Photostrom, so ist die so aufgenommene Kurve mit der zuerst definierten identisch, wie man sofort sieht. In diesem Fall erhält man für verschiedene Photoströme stets die gleiche Kurve φ_λ , genau wie oben für die verschiedenen Energiewerte des eingestrahnten äquinergetischen Spektrums. Wenn diese Proportionalität nicht besteht, so ergibt sich eine Schar von φ_λ -Kurven mit dem Photostrom i als Parameter.

Sind man nun den Durchlaßkoeffizienten eines farbigen Glases nach der 2. Methode bei einem bestimmten Photostrom i , so läßt sich dieser gemessene Durchlaßkoeffizient nach Formel (1) berechnen, wenn man für φ_λ aus der Kurvenschar diejenige mit dem Parameter i einsetzt, und zwar ist diesmal in Zähler und Nenner das gleiche φ_λ einzusetzen, da sowohl im gefilterten als auch im ungefilterten Licht bei dem konstanten Photostrom i gemessen wird.

Da sich bei der Messung für den Durchlaßkoeffizienten kein konstanter Wert ergibt, sondern eine Abhängigkeit vom Photostrom, so bedeutet das nach Formel (1), daß zu verschiedenen Photoströmen tatsächlich verschiedene φ_λ -Kurven gehören, bzw. — was sich inhaltlich mit dieser Aussage deckt — die angenommene Proportionalität zwischen Photostrom und eingestrahelter Lichtintensität nicht besteht.

Sind die φ_λ -Kurven nicht bekannt, so könnte man sich umgekehrt eine φ_λ -Kurve so konstruieren, daß sich mit ihr aus (1) gerade der gemessene Wert D^* ergibt. Freilich läßt sich aus einem einzigen gemessenen Wert das φ_λ aus Gleichung (1) nicht ermitteln. Eine näherungsweise Auswertung des Integrals durch Summation erfordert ja für jeden zu berechnenden Punkt der Kurve φ_λ einen gemessenen Wert von D^* , damit die Zahl der Gleichungen mit der Zahl der Unbekannten übereinstimmt. Um n Punkte der Kurve φ_λ berechnen zu können, müßte man die Durchlaßkoeffizienten von n Farbfiltern messen.

So ist selbst eine näherungsweise Berechnung der Kurve φ_λ aus einer kleinen Anzahl von Farbfiltermessungen nicht möglich. Immerhin läßt sich aber eine Abschätzung über die Änderung der φ_λ -Kurve bei verschiedener Strombelastung des Elementes aus den Messungen der Durchlaßkoeffizienten von drei Farbfiltern durchführen, wie weiter unten gezeigt werden soll.

Experimentelle Methode

Als Beleuchtungsquelle wurde eine Wolfram-Glühlampe (4,35 A, 6 V) von der „Schwarzen“ Temperatur 2680° abs. benutzt.

Es wurden 3 verschiedene Gläser in den Farben rot, grün und blau zu den Messungen verwandt. In verschiedenen Entfernungen von der Lichtquelle wurden die Photoströme von 2 Selen-Photoelementen mit vorgeschaltetem Farbfilter bestimmt. Die Messung der Photoströme wurde mit einem Multiflex-Spiegelgalvanometer mit 1200 Ω Innenwiderstand ausgeführt. Um die Wirkungen verschiedener Außenwiderstände auf die Photoelemente zu untersuchen, wurden außer dieser Meßreihe mit dem Außenwiderstand 1200 Ω noch zwei weitere mit 100 Ω und 10000 Ω Außenwiderstand durchgeführt. Diese Außenwiderstände wurden durch Parallel- und Vorschalten passender Stöpselwiderstände vor das Galvanometer hergestellt.

Es wurden 2 Selen-Photoelemente der Firma Dr. B. Lange untersucht, eines der üblichen Bauart (Element II, 28 mm Durchmesser) und ein Element mit Cd-Zwischenschicht (Element 14/21 28 mm Durchmesser), welches eine gesteigerte Infrarotempfindlichkeit zeigte.

Wie wir später noch sehen werden, ist es zweckmäßig, daß bei der Messung die Photoströme für die einzelnen Farben annähernd gleich sind. Deshalb mußten für die beiden Photoelemente verschiedene Filtergruppen rot, grün und blau gewählt werden, für das Element II die Schott-Filter RG 5/2 mm, VG 9/4 mm und BG 25/2 mm. Für das Element 14/21 wurde das gleiche Blaufilter benutzt, als Grünfilter diente eine Kombination von VG 9/2 mm und VG 6/2 mm und im Roten ein RG 5/2 mm und zur Schwächung ein VG 5/1 mm.

Auswertung der Messergebnisse

Der Durchlaßkoeffizient eines farbigen Glases, gemessen bei dem konstanten Photostrom i , war (2)

$$D^* = \frac{r_2^2}{r_1^2},$$

wenn für beide Messungen die gleiche Lichtquelle unter den gleichen Betriebsbedingungen benutzt wird. Die experimentelle Bestimmung dieses Durchlaßkoeffizienten erfordert eine beträchtliche Länge der optischen Bank, sowie eine über die ganze Länge gewährleistete Gültigkeit des Entfernungsgesetzes. Daher erscheint es zweckmäßig, nicht die Durchlaßkoeffizienten selbst, sondern deren Verhältnisse zu bestimmen. Hierbei unterbleibt die Messung im ungefilterten Licht, da das Verhältnis des Durchlaßkoeffizienten D_R^* des Rotglases zum Durchlaßkoeffizienten D_G^* des Grünglases z. B. gegeben ist als

$$\frac{D_R^*}{D_G^*} = \frac{r_R^2}{r_G^2},$$

wenn r_R und r_G die Abstände des Photoelementes von der Lichtquelle mit eingeschaltetem Rot- bzw. Grünfilter bedeuten und die übrigen Bedingungen wie oben erfüllt sind.

Wählt man die Farbfilter so, daß die Entfernungen r_R und r_G nur wenig voneinander verschieden sind, so werden die Photoströme i_R und i_G , die man erhält, wenn man das Photoelement durch das Rot- bzw. Grünfilter in konstanter — zwischen r_R und r_G gelegenen — Entfernung von der Lichtquelle bestrahlt, eben-

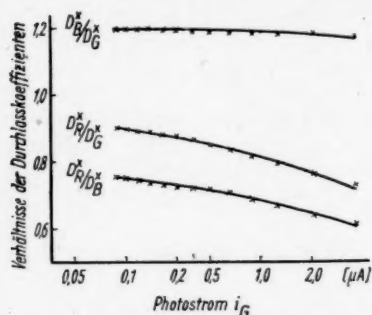


Abb. 2. Abhängigkeit der Verhältnisse des Durchlasskoeffizienten vom Photostrom i_G . Photoelement II, $R_A = 1200 \Omega$

Benutzt man nun 3 Farbgläser rot, grün und blau, die so gewählt sind, daß die in fester Entfernung von der Lichtquelle gemessenen Photoströme i_R , i_G , i_B nur wenig voneinander verschieden sind, so kann man die Verhältnisse der Durchlasskoeffizienten bestimmen als

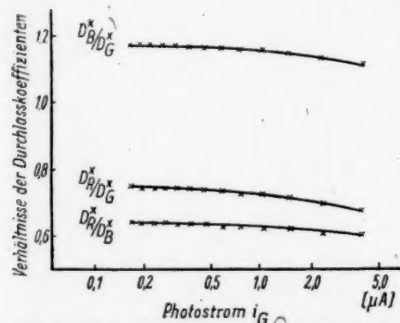


Abb. 3. Abhängigkeit der Verhältnisse des Durchlasskoeffizienten vom Photostrom i_G . Photoelement 14/21 mit Cd-Zwischenschicht, $R_A = 1200 \Omega$

strom i_G (Strom mit eingeschaltetem Grünfilter) auf. Die Kurven für den Außenwiderstand 1200Ω sind in den Abb. 2 und 3 dargestellt. Als Abszisse wurde i_G in logarithmischem Maßstab gewählt.

Um zu prüfen, ob die aus den Photoströmen bei konstantem Abstand berechneten Verhältnisse der Durchlasskoeffizienten wirklich mit den bei konstantem Strom durch die Entfernungsquadrate bestimmten Verhältnissen der Durch-

falls nur wenig voneinander verschieden sein. Das Verhältnis i_R/i_G wird dann annähernd gleich $r_R^2/r_G^2 = D_R^x/D_G^x$ sein. Wie aus dem oben Gesagten hervorgeht, wäre diese Beziehung genau erfüllt, falls zwischen Photostrom und eingestrahler Lichtintensität strenge Proportionalität bestünde. Wegen der geringen Entfernungs- bzw. Stromunterschiede der beiden Messungen ist für diesen kleinen Bereich die Proportionalität genügend genau erfüllt. Die oben erwähnte Beziehung ist also im Rahmen der Meßgenauigkeit als gültig anzusehen. Das Verhältnis D_R^x/D_G^x ist dann auf einen mittleren Strom zwischen i_R und i_G zu beziehen.

$$D_R^x/D_G^x = i_R/i_G$$

$$D_R^x/D_B^x = i_R/i_B \quad (3)$$

$$D_B^x/D_G^x = i_B/i_G.$$

Wählt man die Filter noch so, daß i_G ungefähr in der Mitte zwischen i_R und i_B liegt, so kann man die so bestimmten Verhältnisse auf den mittleren Strom i_G beziehen.

Wir berechnen aus den Messungen die Verhältnisse der Durchlasskoeffizienten nach Gl. (3) für verschiedene Entfernungen von der Lichtquelle und tragen sie — wie oben angegeben — über dem Photo-

laßkoeffizienten übereinstimmen, wurden diese für einen konstanten Photostrom berechnet. Die Kurven stimmen dem Verlauf und der Größenordnung nach mit den Abb. 2 und 3 überein.

Aus allen Kurven geht die Abhängigkeit der Verhältnisse der Durchlaßkoeffizienten und somit der Durchlaßkoeffizienten selbst von der Strombelastung deutlich hervor.

Diese Abhängigkeit des Durchlaßkoeffizienten vom Photostrom läßt sich, wie wir oben gesehen haben, nach Gl. (1) als eine Änderung der spektralen Empfindlichkeit mit der Strombelastung deuten. Ein Überblick über die Größenordnung dieser Änderung soll hier gegeben werden.

Nach Formel (1) ergab sich z. B. für das Verhältnis $D_R^{(i)}/D_G^{(i)}$, gemessen bei dem Photostrom i

$$\frac{D_R^{(i)}}{D_G^{(i)}} = \frac{\int_0^\infty \tau_{\lambda R} E_\lambda \varphi_\lambda^{(i)} d\lambda}{\int_0^\infty \tau_{\lambda G} E_\lambda \varphi_\lambda^{(i)} d\lambda}, \quad (1a)$$

wobei $\tau_{\lambda R}$ bzw. $\tau_{\lambda G}$ die spektrale Durchlässigkeit des Rot- bzw. Grünlasses, E_λ die relative spektrale Energieverteilung der Lampe und $\varphi_\lambda^{(i)}$ die spektrale Empfindlichkeit des Photoelementes für den konstanten Photostrom i (jeweils für die Wellenlänge λ) bedeuten, wie sie oben definiert wurde. Kennt man für einen bestimmten Photostrom i_1 die Kurve $\varphi_\lambda^{(i_1)}$, so kann man diese in Formel (1a) einsetzen und prüfen, ob das berechnete Verhältnis der Durchlaßkoeffizienten mit dem gemessenen übereinstimmt. (Für $D_R^{(i)}/D_G^{(i)}$ ergibt sich aus (1) eine (1a) entsprechende Formel.) Ist dies mit genügender Genauigkeit erfüllt, so kann man für einen Photostrom i_2 , der seiner Größe nach von i_1 nicht allzusehr abweicht, durch systematische probeweise Änderung der Kurve $\varphi_\lambda^{(i_1)}$ eine Kurve $\varphi_\lambda^{(i_2)}$ konstruieren derart, daß sich mit ihr aus (1a) das für den Photostrom i_2 gemessene Verhältnis der Durchlaßkoeffizienten ergibt.

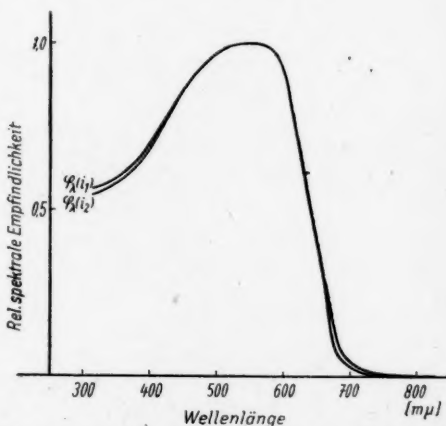


Abb. 4. Die relative spektrale Empfindlichkeit des Photoelementes II, $R_A = 1200 \Omega$ bei verschiedenen Photostromen. $i_1 = 0,09 \mu A$, $i_2 = 2,30 \mu A$

In unserem Fall liegt keine für konstanten Strom aufgenommene Kurve der spektralen Empfindlichkeit vor. Man kann aber statt einer solchen eine Kurve φ_λ benutzen, die nach dem üblichen Verfahren durch die Änderung des Photostroms bei eingestrahlttem äquinergetischen Spektrum bestimmt ist, sofern die Energie

des Spektrums nicht allzu groß ist. Abb. 4 und 5 stellen die auf diese Weise erhaltenen φ_λ -Kurven für die beiden Photoelemente dar.

Man erkennt, daß die Änderungen hier nur gering sind, obwohl für beide Photoelemente der Photostrom im 2. Fall etwa das 25fache des Anfangswertes beträgt. Freilich sind die bei unseren Messungen auftretenden Photoströme nur sehr klein. Geht man zu wesentlich höheren Photoströmen etwa bis zu $500 \mu\text{A}$ über, wie sie bei stärkeren Beleuchtungen in der Praxis vorkommen, so werden die Änderungen der φ_λ -Kurve wahrscheinlich größer sein. Aus den bisher vorliegenden Messungen in diesem Strombereich, die von Dresler durchgeführt wurden, läßt sich eine hinreichend genaue Aussage über die Änderung der φ_λ -Kurve nicht machen, da die

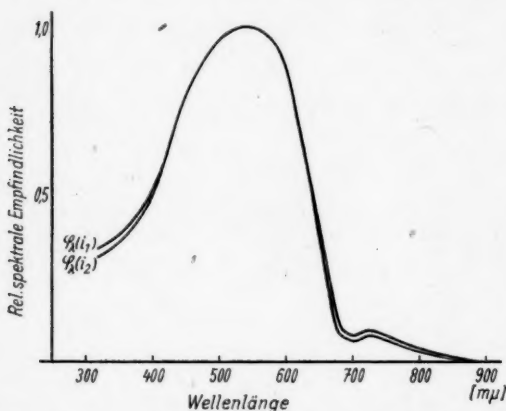


Abb. 5. Die relative spektrale Empfindlichkeit des Photoelementes 14/21 mit Cd-Zwischenschicht, $R_A = 1200 \Omega$, bei verschiedenen Photoströmen. $i_1 = 0,17 \mu\text{A}$, $i_2 = 4,02 \mu\text{A}$

zogen auf die Empfindlichkeit der Photoelemente gewählt. Diese ergibt sich folgendermaßen. Wir setzen willkürlich die in der Entfernung $r_0 = 70 \text{ cm}$ für grünes Licht herrschende Beleuchtungsstärke E_{0g} gleich 1 und erhalten somit für grünes Licht die relative Beleuchtungsstärke in der Entfernung r

$$E_G = \frac{70^2}{r^2}.$$

Wir messen nun die Beleuchtungsstärken im roten und blauen Licht für die Entfernung r ebenfalls in unserer Beleuchtungsstärkeeinheit.

$$E_R = \bar{E}_{0R} \cdot \frac{70^2}{r^2} \quad E_B = \bar{E}_{0B} \cdot \frac{70^2}{r^2}.$$

\bar{E}_{0R} und \bar{E}_{0B} sind die in der Entfernung r_0 herrschenden Beleuchtungsstärken im roten und blauen Licht, bezogen auf unsere willkürlich festgelegte Einheit im grünen Licht. Diese Größen sind ja die Verhältnisse der Beleuchtungsstärken in der Entfernung r_0 und werden mit dem Photoelement als das Verhältnis der Durch-

dort verwandten Filter einen zu großen Durchlaßbereich besitzen.

Um die für die einzelnen Farben verschiedene Nichtproportionalität zwischen Photostrom und eingestrahelter Lichtintensität, die ja die Ursache für das Auftreten verschiedener φ_λ -Kurven ist, erkennen zu können, ist in den Abb. 6 und 7 die relative Empfindlichkeit der beiden Photoelemente für rotes, grünes und blaues Licht bei den 3 Außenwiderständen 100Ω , 1200Ω , 10000Ω aufgetragen. Als Abszisse wurde die relative Beleuchtungsstärke $E^{(2)}$ be-

laßkoeffizienten bestimmt. Wir wollen die auf die spektrale Empfindlichkeit des Photoelementes bezogenen Beleuchtungsstärken in der Entfernung r mit dem oberen Index z versehen.

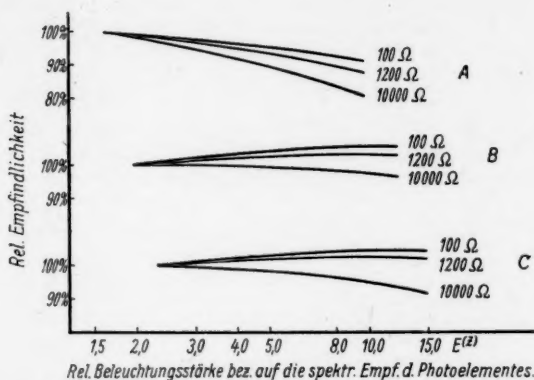


Abb. 6. Abhängigkeit der relativen Empfindlichkeit von der relativen Beleuchtungsstärke $E^{(z)}$ für die Außenwiderstände 100 Ω , 1200 Ω und 10000 Ω bei Bestrahlung auf rotem (A), grünem (B) und blauem (C) Licht. Photoelement II

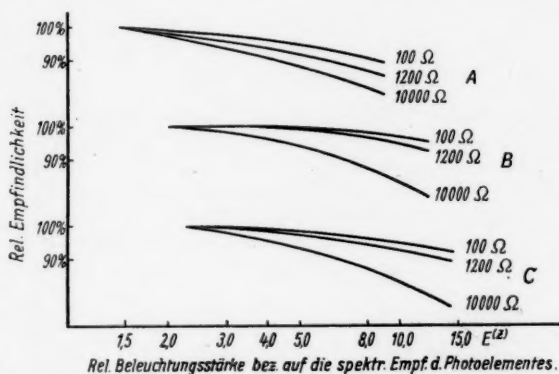


Abb. 7. Abhängigkeit der relativen Empfindlichkeit von der relativen Beleuchtungsstärke $E^{(z)}$ für die Außenwiderstände 100 Ω , 1200 Ω und 10000 Ω bei Bestrahlung mit rotem (A), grünem (B) und blauem (C) Licht. Photoelement 14/21 mit Cd-Zwischenschicht

Wir bilden die Quotienten $i_R/E_R^{(z)}$, $i_G/E_G^{(z)}$ und $i_B/E_B^{(z)}$. Dies ist die sogenannte „Empfindlichkeit“ des Photoelementes für rotes, grünes und blaues Licht.

Ist der Photostrom der Beleuchtungsstärke proportional, so ergibt sich, wenn wir die Empfindlichkeit über der Beleuchtungsstärke auftragen, eine Parallele

zur Abszissenachse. Die Abweichung der Kurven von diesem idealen Fall gibt uns ein Maß für die vorhandene Nichtproportionalität.

In den Abb. 6 und 7 sind die Kurven der Empfindlichkeit für die 3 Außenwiderstände für jede einzelne Farbe auf einen gemeinsamen Anfangspunkt bezogen, dessen Ordinate für alle Widerstände gemeinsam gleich 100 gesetzt wurde. Man erkennt also hier direkt, daß die prozentuale Abweichung vom idealen Verlauf mit wachsender Beleuchtungsstärke für höhere Widerstände in jeder einzelnen Farbe größer ist als für kleinere. Alle Kurven zeigen schließlich für größere Beleuchtungsstärken eine fallende Tendenz, d. h. der Photostrom steigt langsamer an als die zugehörige Beleuchtungsstärke. Dieses Verhalten macht sich mit wachsendem Außenwiderstand immer stärker bemerkbar.

Für die Hilfe bei der experimentellen Durchführung meiner Arbeit danke ich Herrn Prof. Dr. Kienle und den Mitarbeitern des Astrophysikalischen Observatoriums Potsdam. Herrn Prof. Dr. Wempe danke ich für seine freundliche Unterstützung und manchen wertvollen Rat.

Potsdam, Astrophysikalisches Institut.

(Bei der Redaktion eingegangen am 1. November 1949.)

Zum Problem der Maxwellschen Spannungen

Von A. Sommerfeld und F. Bopp

Inhaltsübersicht

Um die verschiedenen, in letzter Zeit vorgeschlagenen Ausdrücke für die magnetostatische Kraftdichte \vec{j} prüfen zu können, fügen wir den Maxwellschen Gleichungen das Postulat hinzu, daß sich \vec{j} als Divergenz eines symmetrischen Spannungstensors berechnen lassen solle, nicht nur außerhalb des Magneten, sondern auch im Innern desselben. Der klassische Maxwell-Tensor führt in symmetrisierter Form weder auf den Ansatz $\vec{j} = \text{div } \mathfrak{S} \cdot \mathfrak{B}$ noch auf den Ansatz $\vec{j} = \mu \text{div } \mathfrak{S} \cdot \mathfrak{S}$, sondern auf die kompliziertere Gl. (11). Neben den Maxwellschen stellen wir einen „modifizierten Spannungstensor“ unter Hinweis auf eine kürzlich erschienene Darstellung von L. Kneissler. Die zu diesem Tensor gehörende Kraftdichte wird durch die Vorstellung der Ampere-schen Molekularströme nahegelegt. Das aus ihm folgende Drehmoment verhält sich zu dem aus dem Maxwell-Tensor folgenden wie $\mu:\mu_0$ (μ = Permeabilität des Außenraumes). Durch Messung des Drehmomentes kann man also zwischen diesen beiden Tensoren und den ihnen entsprechenden Kraftdichten entscheiden. Es wird aber betont, daß von unserem rein phänomenologischen Standpunkt aus die Form des Spannungstensors nicht eindeutig bestimmt ist und daß auch die Messung des Drehmomentes noch keinen eindeutigen Aufschluß über die zugrunde liegende Kraftdichte liefert.

§ 1. Vorbemerkungen

Die Relativitätstheorie faßt die elektrische Feldstärke \mathfrak{E} und die magnetische Induktion \mathfrak{B} zu einem Sechservektor (antisymmetrischen Tensor) zusammen, ebenso die elektrische Verschiebung \mathfrak{D} und die sog. magnetische Feldstärke \mathfrak{H} . Was die Relativitätstheorie zusammengefügt hat, soll der Mensch nicht trennen. \mathfrak{B} hat danach denselben Charakter wie \mathfrak{E} , nämlich den einer Feldstärke (Intensitätsgröße), \mathfrak{H} ist ebenso wie \mathfrak{D} eine Quantitätsgröße. Die Unterscheidung zwischen \mathfrak{B} und \mathfrak{H} wird durch Dimensionsbetrachtungen in bekannter Weise unterstrichen.

In der phänomenologischen Darstellung der Magnetostatik ist

$$\text{div } \mathfrak{B} = 0, \quad \text{rot } \mathfrak{H} = 0, \quad \mathfrak{H} = -\text{grad } \varphi \quad (1)$$

mit φ als magnetischem Potential. Es gilt daher

$$\Delta \varphi = -\text{div } \mathfrak{H}, \quad 4\pi \varphi = \int \frac{\text{div } \mathfrak{H}}{r} d\tau;$$

$\text{div } \mathfrak{H}$ spielt für die Feldberechnung sicherlich die Rolle der magnetischen Dichte. Ob $\text{div } \mathfrak{H}$ dieselbe Bedeutung für die Berechnung der Kraftwirkungen habe, soll in Folgendem untersucht werden. Zunächst liegt es von hieraus nahe, als Ausdruck für die Kraftdichte anzusetzen

$$\vec{j} = \text{div } \mathfrak{H} \cdot \mathfrak{B}. \quad (2)$$

Dieser Ansatz wurde in der ursprünglichen Fassung der nachfolgend mitgeteilten Rechnung von Sommerfeld und Ramberg durchgeführt, ist aber aus den in dieser Note zu besprechenden Gründen fallen gelassen worden.

Wir vergleichen den Ansatz (2) mit dem jetzt allgemein anerkannten Ausdruck für die Lorentzsche Kraftdichte

$$\vec{f} = e \vec{E} + [\vec{j}, \vec{B}], \quad (3)$$

wo $\vec{j} = \rho \vec{v}$ die Dichte des elektrischen Stromes ist. In dieser Form, mit \vec{B} statt des in der älteren Literatur üblichen \vec{H} geschrieben, folgt (3) aus der Relativitätstheorie.

Gehen wir zur Magnetostatik $\vec{E} = 0$ über und identifizieren wir die in (3) vorkommende Stromdichte ausschließlich mit dem Magnetisierungsstrom $\text{rot } \vec{M}$, so kommen wir zu dem Ansatz

$$\vec{f} = \frac{1}{\mu_0} [\text{rot } \vec{B}, \vec{B}], \quad (4)$$

der formal an (3) anschließt, aber mit (2) im Widerspruch steht. Er liegt einer Arbeit von Hallén¹⁾ zugrunde und wird in einer Broschüre von Kneissler²⁾ eingehend diskutiert.

Wir sehen hieraus, daß die Kraftdichte durch die Maxwell'schen Gln. (1) keineswegs eindeutig bestimmt ist. Ein Kriterium zur Beurteilung der verschiedenen Möglichkeiten, das durchaus im Faraday-Maxwell'schen Gedankenkreis liegt, liefert das weitere Postulat, daß die Kraftdichte sich als Divergenz eines Spannungstensors darstellen lasse. Dieser Spannungstensor braucht „im Allgemeinen“ nicht mit dem Maxwell'schen überein zu stimmen. Der „allgemeine Fall“ möge charakterisiert sein durch

$$\vec{B} = \mu_0 (\vec{S} + \vec{M}), \quad (5)$$

unter \vec{M} die Magnetisierungs-dichte (gewöhnlich mit $4\pi\vec{S}$ bezeichnet), unter μ_0 den konstanten Dimensionsfaktor (Vakuumpermeabilität) verstanden. \vec{M} setzt sich aus einem von der Suszeptibilität des Materials herrührenden und einem permanenten Bestandteil zusammen:

$$\vec{M} = \kappa \vec{S} + \vec{M}_0 \quad (6)$$

\vec{M}_0 und κ , also auch \vec{M} sind als ortsabhängige Größen anzusehen, eventuell als unstetige Funktionen, die in verschiedenen Gebieten verschiedene konstante Werte haben.

Da es in der Magnetostatik wesentlich auf die Berechnung des auf einen Körper wirkenden Drehmomentes ankommt, müssen wir das weitere Postulat hinzufügen, daß der Spannungstensor symmetrisch sein solle: $T_{ik} = T_{ki}$. Nur dann ist nämlich das Drehmoment \vec{N} aus der Kraftdichte \vec{f} nach der anschaulichen Formel

$$\vec{N} = \int [\vec{r}, \vec{f}] d\tau \quad (7)$$

zu berechnen, während bei unsymmetrischen T_{ik} Glieder hinzutreten würden, die von der Differenz $T_{ik} - T_{ki}$ abhängen und die elementare Berechnung (7) von \vec{N} illusorisch machen würden.

¹⁾ E. Hallén, Kungl. Tekn. Högsk. Handl. Nr. 6, 1947.

²⁾ L. Kneissler, Die Maxwell'sche Theorie in veränderter Formulierung, Springer, Wien 1949.

§ 2. Der symmetrisierte Maxwell'sche Spannungstensor

Wir lassen, wie in der Tensorrechnung üblich, die Summationszeichen bei Gliedern mit zweimal vorkommendem Index fort und schreiben abkürzend

$\partial/\partial_i, \partial/\partial_k \dots$ statt $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ ($i, k \dots$ durchlaufen die Koordinaten x, y, z).

Offenbar ist der symmetrische Tensor

$$T_{ik} = \frac{1}{2} (H_i B_k + H_k B_i - \delta_{ik} H_j B_j) \quad (8)$$

identisch mit dem klassischen Maxwell'schen Tensor, wenn Proportionalität zwischen B und H besteht; wenn dagegen die allgemeine Beziehung (5) gilt, kann (8) als eine Verallgemeinerung des Maxwell-Tensors angesehen werden. Die zugehörige Kraftdichte f_i in Richtung der i -Koordinate berechnet sich wie in der Elastizitätstheorie als Tensor-Divergenz:

$$f_i = \frac{\partial T_{ik}}{\partial k}. \quad (8a)$$

Man erhält

$$2 f_i = H_i \frac{\partial B_k}{\partial k} + B_k \frac{\partial H_i}{\partial k} + H_k \frac{\partial B_i}{\partial k} + B_i \frac{\partial H_k}{\partial k} - \frac{\partial H_j B_j}{\partial i}. \quad (9)$$

Das erste Glied rechts verschwindet wegen $\text{div } \mathfrak{B} = 0$, das zweite ist wegen des Verschwindens von \mathfrak{S} gleich

$$B_k \frac{\partial H_k}{\partial i} = \frac{\partial B_k H_k}{\partial i} - H_k \frac{\partial B_k}{\partial i}.$$

Hier hebt sich das erste Glied rechts gegen das letzte Glied von (9), das zweite vereinigen wir mit dem dritten Glied von (9). Wir erhalten so:

$$2 f_i = H_k \left(\frac{\partial B_i}{\partial k} - \frac{\partial B_k}{\partial i} \right) + B_i \frac{\partial H_k}{\partial k}. \quad (10)$$

Dies ist in vektorieller Schreibweise gleich

$$2 \mathfrak{f} = [\text{rot } \mathfrak{B}, \mathfrak{S}] + \text{div } \mathfrak{S} \cdot \mathfrak{B}. \quad (11)$$

Unser Ergebnis bestätigt also nicht den Ansatz (2). Vielmehr ergibt sich \mathfrak{f} als arithmetisches Mittel zweier Terme, deren einer mit (2) übereinstimmt, während der andere dem Ansatz (4) ähnelt und von (2) im allgemeinen verschieden ist.

Ebensowenig bestätigt unser Ergebnis den Ansatz

$$\mathfrak{f} = \mu \text{div } \mathfrak{S} \cdot \mathfrak{S}. \quad (12)$$

Herr Döring³⁾ zeigt, daß dieser Ansatz, in die vorstehende Gl. (7) eingesetzt, denselben Wert des Drehmomentes \mathfrak{M} liefert wie die Berechnung aus dem Max-

³⁾ W. Döring, Ann. Physik (6) 6, 69 (1949), speziell Gl. (27) und (28). Herr Döring weist uns freundlichst darauf hin, daß Gl. (12) als Ansatz für die Kraftdichte explicite nicht bei ihm vorkommt. Wir stimmen mit ihm auch darin überein, daß Gl. (12) hierfür schon darum nicht in Betracht kommt, weil das äußere μ im Innern des Magneten keine physikalische Bedeutung hat. Andererseits bemerken wir, daß wir uns auch in Widerspruch mit Gl. (11) setzen würden, wenn wir den Faktor μ in (12) als reversible Permeabilität des Magneten [in unserer Bezeichnungsweise $\mu_0(1 + \kappa)$] auffassen wollten.

wellschen Spannungstensor. Es handelt sich bei Döring, ebenso wie in der folgenden Note, um das von einem äußeren Felde auf einen magnetisierten Körper ausgeübte Drehmoment, wobei sich der Körper in einer Umgebung von der konstanten Permeabilität μ befindet.

§ 3. Der modifizierte Spannungstensor

Als solchen wollen wir betrachten

$$T_{ik} = \frac{1}{\mu_0} (B_i B_k - \frac{1}{2} B_j^2 \delta_{ik}). \quad (13)$$

Er stimmt mit dem von L. Kneissler l. c., Gl. III (38) eingeführten Tensor überein (abgesehen von der unwesentlichen Wahl der Einheiten). Indem wir durch Divergenzbildung zur Kraftdichte f_i übergehen, erhalten wir

$$\mu_0 f_i = \frac{\partial B_i}{\partial k} B_k + \frac{\partial B_k}{\partial k} B_i - B_j \frac{\partial B_j}{\partial i}. \quad (14)$$

Hier verschwindet der zweite Term wegen $\text{div } \mathfrak{B} = 0$. Im dritten können wir den Summationsindex j mit k vertauschen. Dadurch vereinfacht sich (14) zu

$$\mu_0 f_i = \left(\frac{\partial B_i}{\partial k} - \frac{\partial B_k}{\partial i} \right) B_k. \quad (15)$$

Dies ist mit dem Ansatz (4) identisch, wenn wir wie in (11) zur vektoriellen Schreibweise übergehen. Über den Zusammenhang von \mathfrak{B} und \mathfrak{S} ist dabei nichts ausgesagt; er kann z. B. in der allgemeinen Form der Gl. (5) und (6) angenommen werden bei beliebiger räumlicher Verteilung von \mathfrak{M}_0 und κ .

Zur praktischen Berechnung des Drehmomentes wird man in dem von Döring behandelten Falle nicht von dem Raumintegral (7) ausgehen, sondern man wird das Flächenintegral der Momente des Spannungstensors bilden und dieses über eine äußere Hüllfläche des Körpers erstrecken. Man kann dann die einfache Beziehung $\mathfrak{B} = \mu \mathfrak{S}$ benutzen und braucht das Verhalten auf der Grenzfläche und im Körperinnern nicht zu untersuchen. Bei dieser Berechnung liefert unser modifizierter Spannungstensor bis auf den konstanten Faktor μ/μ_0 offenbar dasselbe wie der Maxwell-Tensor. In der Tat gilt ja auf der Hüllfläche nach (13) bzw. (8)

$$T_{ik} = \frac{\mu^2}{\mu_0} (H_i H_k - \frac{1}{2} H_j^2 \delta_{ik})$$

bzw.

$$T_{ik} = \mu (H_i H_k - \frac{1}{2} H_j^2 \delta_{ik}).$$

In demselben Verhältnis wie die beiden Spannungstensoren stehen auch die aus beiden folgenden Drehmomente. Von da aus dürfte sich die Möglichkeit ergeben, experimentell zwischen den beiden Spannungstensoren und den zugehörigen Kraftdichten (4) und (11) zu unterscheiden. Die beiden Ansätze (2) und (12) scheinen uns bei dieser Prüfung aus den oben genannten Gründen auszuschneiden.

Die Möglichkeit verschiedenartiger Ansätze für den Spannungstensor zeigt, daß die vorangestellten Postulate, Existenz und Symmetrie eines Spannungs-

tensors, keineswegs ausreichen, um das pondermotorische Verhalten magnetisierter Körper eindeutig zu bestimmen. Dazu wären weitere Erfahrungen nötig, nicht nur Messungen des Drehmomentes, sondern womöglich solche der lokalen Kraftdichte im Innern oder an der Oberfläche des Magneten.

Man kann die theoretische Seite des Problems etwas vereinfachen, wenn man statt des Tensors T_{ik} den Skalar W der Energiedichte untersucht. Es läßt sich nämlich T_{ik} nach dem Prinzip der virtuellen Arbeit aus W ableiten, wenn man noch fordert, daß der Spannungstensor und die Energiedichte im Vakuum gleich den Maxwell'schen Ausdrücken sein sollen. Z. B. folgt der symmetrisierte Maxwell-Tensor (8) aus $W = \frac{1}{2} \mathfrak{E} \cdot \mathfrak{B}$ und der modifizierte Tensor (13) aus $W = \frac{1}{2} \mu_0 \cdot \mathfrak{B}^2$ durch ähnliche makrophysikalische Gedankenexperimente, wie sie Becker⁴⁾ für den elektrischen Fall diskutiert hat.

Die atomistisch verfeinerte Durchführung der Lorentz'schen Ansätze für die Elektronenstatistik im Innern eines Magneten müßte zu einem bestimmten Ausdruck für W und daher auch zu einem bestimmten Tensor T_{ik} führen. In der folgenden Note von Sommerfeld und Ramberg wird mit allem Vorbehalt der modifizierte Spannungstensor (13) zugrunde gelegt werden.

⁴⁾ R. Becker, Theorie der Elektrizität I § 41, 42; Leipzig 1949 (auch 1933).

München, Institut für theoretische Physik.

(Bei der Redaktion eingegangen am 20. Mai 1950.)

Das Drehmoment eines permanenten Magneten im Felde eines permeablen Mediums

Von A. Sommerfeld und E. Ramberg¹⁾

Inhaltsübersicht

§ 1. Lösung der Randwertaufgabe: Das verlängerte, homogen magnetisierte Ellipsoid in einem homogenen äußeren Felde beliebiger Richtung. — § 2. Berechnung des Drehmomentes aus einem modifizierten Spannungstensor. — § 3. Ausführung der Integrationen. — § 4. Der Grenzfall des Stabmagneten. — § 5. Das Rotationsellipsoid beliebiger Form, insbesondere im Falle $\mu = \mu_0$.

Da dieses Problem neuerdings in der Literatur mehrfach²⁾ diskutiert worden ist, dürfte eine ins Einzelne gehende mathematische Behandlung desselben erwünscht sein. Die Vorbedingung derselben ist natürlich die Lösung der betreffenden Randwertaufgabe. Da sie für den zylindrischen Stab auf große Schwierigkeiten stößt, ersetzen wir wie üblich den Stab durch ein verlängertes Rotationsellipsoid. Auf die Gefahr hin, daß die in Rede stehende Randwertaufgabe schon sonst behandelt³⁾ worden ist, schreiben wir ihre Lösung in § 1 an. Es sei, wie üblich $a, b =$ große und kleine Achse des Rotationsellipsoides, $c = \sqrt{a^2 - b^2} =$ lineare Exzentrizität. Der Grenzfall $b \ll a$ vermittelt den Übergang zum Stabmagneten.

In § 2 und ff. wird das Drehmoment berechnet. Wir gehen dabei aus von dem in der vorangehenden Note (Sommerfeld-Bopp) definierten, modifizierten Spannungstensor, nicht von einem Elementargesetz für die Kraftdichte, welches nach den Ausführungen der vorangehenden Note erst durch die Wahl des Spannungstensors festgelegt wird. Übrigens unterscheidet sich, wie dort gezeigt, das so berechnete Drehmoment von dem aus dem Maxwell'schen Tensor folgenden nur um den Faktor μ/μ_0 . Auf diese Note verweisen wir auch wegen der grundsätzlichen Frage, die den Anlaß zu der nachfolgenden Untersuchung gegeben hat, ob nämlich \mathfrak{B} oder \mathfrak{H} als dynamischer Vektor aufzufassen sei.

§ 1. Lösung der Randwertaufgabe

Das Ellipsoid sei homogen magnetisiert nach der großen Achse, die wir zur z -Achse machen; \mathfrak{M} sei das magnetische Moment der Volumeneinheit. Wir sehen

¹⁾ Princeton, N. J., Laboratorium der R. C. A.

²⁾ H. Diesselhorst, Ann. Physik **3**, 11 (1948); W. Döring, Ann. Physik **6**, 69 (1949).

³⁾ Herr Diesselhorst weist uns auf eine Arbeit von Leigh Page, Physic. Rev. **44**, 113 (1933) hin, in der ein Spezialfall unseres Problems, nämlich das „Eigenfeld“ des homogen magnetisierten Ellipsoids behandelt wird, bei Absehung von einem äußeren „Fremdfeld“. Die Frage nach dem Drehmoment kommt bei dieser Spezialisierung des Problems natürlich in Fortfall.

\mathfrak{M} als fest gegebene Größe an (sehr hartes magnetisches Material! Suszeptibilität $\kappa = 0$!). Die Richtung des äußeren Feldes liege in der x - z -Ebene, wir charakterisieren es durch sein Potential

$$\psi_0 = -(\alpha x + \gamma z) H \quad (1)$$

μ_0 und μ seien die Permeabilitäten⁴⁾ des Ellipsoids und seiner Umgebung. Im Äußeren bzw. Inneren ist

$$\mathfrak{B} = \mu \mathfrak{H} \text{ bzw. } \mathfrak{B} = \mu_0 (\mathfrak{H} + \mathfrak{M}) \quad (2)$$

\mathfrak{H} leitet sich aus einem Potential ab gemäß

$$\mathfrak{H} = -\text{grad } \psi, \quad \psi = \begin{cases} \psi_{\text{I}}, & \text{Inneres} \\ \psi_{\text{II}}, & \text{Äußeres,} \end{cases} \quad (3)$$

ψ ist eine stetige Funktion des Ortes, die überall der Laplaceschen Gleichung

$$\Delta \psi = 0 \quad (4)$$

genügt; nur an der Oberfläche des Ellipsoids findet (induzierter Magnetismus!) ein Sprung im normalen Gradienten von ψ statt. Im Unendlichen gilt die Bedingung

$$\psi \rightarrow \psi_0. \quad (4a)$$

Wir führen die elliptischen Koordinaten u, v, φ ein durch

$$\begin{aligned} x &= c \, \text{Sin } u \sin v \cos \varphi \\ y &= c \, \text{Sin } u \sin v \sin \varphi \\ z &= c \, \text{Cos } u \cos v. \end{aligned} \quad (5)$$

Die Oberfläche des Ellipsoids sei $u = u_0$. Hier wird für

$$\begin{aligned} v = 0: & \quad x = y = 0, \quad z = a, \quad \text{Cos } u_0 = a/c, \\ v = \frac{\pi}{2}: & \quad z = 0, \quad \sqrt{x^2 + y^2} = c \, \text{Sin } u_0, \quad \text{Sin } u_0 = b/c. \end{aligned} \quad (5a)$$

Das Linienelement ist gegeben durch

$$ds^2 = c^2 (\text{Cos}^2 u - \cos^2 v) (du^2 + dv^2) + c^2 \text{Sin}^2 u \sin^2 v d\varphi^2. \quad (6)$$

Daraus folgt

$$\begin{aligned} ds_u &= c \sqrt{\text{Cos}^2 u - \cos^2 v} du, \\ ds_v &= c \sqrt{\text{Cos}^2 u - \cos^2 v} dv, \\ ds_\varphi &= c \text{Sin } u \sin v d\varphi. \end{aligned} \quad (6a)$$

Gl. (4) lautet in den u, v, φ

$$\frac{\partial}{\partial u} \left(\text{Sin } u \sin v \frac{\partial \psi}{\partial u} \right) + \frac{\partial}{\partial v} \left(\text{Sin } u \sin v \frac{\partial \psi}{\partial v} \right) + \frac{\text{Cos}^2 u - \cos^2 v}{\text{Sin } u \sin v} \frac{\partial^2 \psi}{\partial \varphi^2} = 0. \quad (7)$$

Natürlich wird sie befriedigt durch die schon in (1) enthaltenen Lösungen

$$\psi = \frac{z}{c} = \text{Cos } u \cos v \text{ und } \psi = \frac{x}{c} = \text{Sin } u \sin v \cos \varphi. \quad (7a)$$

⁴⁾ In unserer Bezeichnungsweise hat μ dieselbe Dimension wie μ_0 , bedeutet also nicht „relative Permeabilität gegen Vakuum“.

Wir suchen je eine zweite partikuläre Lösung desselben Charakters in v und φ . Wir nennen sie

$$\psi = f(u) \cos v \text{ und } \varphi = g(u) \sin v \cos \varphi. \quad (7b)$$

f und g genügen nach (7) den Diff. Gln.

$$\begin{aligned} f'' + \frac{\mathfrak{Cof} u}{\mathfrak{Sin} u} f' - 2f &= 0 \\ g'' + \frac{\mathfrak{Cof} u}{\mathfrak{Sin} u} g' - \left(1 + \frac{\mathfrak{Cof}^2 u}{\mathfrak{Sin}^2 u}\right) g &= 0. \end{aligned} \quad (7c)$$

Wir finden:

$$\begin{aligned} f(u) &= 1 - \mathfrak{Cof} u \log \frac{\mathfrak{Cof} u + 1}{\mathfrak{Sin} u}, \\ g(u) &= \frac{\mathfrak{Cof} u}{\mathfrak{Sin} u} - \mathfrak{Sin} u \log \frac{\mathfrak{Cof} u + 1}{\mathfrak{Sin} u} \end{aligned} \quad (8)$$

(log bedeutet hier und im Folgenden den natürlichen Logarithmus).

Aus (7a) und (7b) setzt sich die allgemeine Lösung des gesuchten Charakters zusammen. Wir müssen aber die zusätzlichen Lösungen (7b) im Inneren ausschließen, weil $f(u)$ und $g(u)$ für $u = 0$ unendlich werden. Unsere allgemeine Lösung des Problems lautet⁵⁾:

$$\begin{aligned} \psi_{\text{I}} &= A \mathfrak{Cof} u \cos v + B \mathfrak{Sin} u \sin v \cos \varphi, \\ \psi_{\text{II}} &= \psi_0 + C f(u) \cos v + D g(u) \sin v \cos \varphi. \end{aligned} \quad (9)$$

Für große Werte von u , wo $\mathfrak{Cof} u = \mathfrak{Sin} u = e^u$ wird, folgt aus (8) $f(u) = g(u) = 1 - e^u \log(1 - e^{-u}) = 1 - (1 - \frac{1}{2}e^{-u} + \dots)$, woraus man sieht, daß unsere Bedingung (4a) durch (9) von selbst erfüllt wird.

(9) enthält vier willkürliche Konstanten A, B, C, D , welche sich durch zwei Paare von Grenzbedingungen für $u = u_0$ bestimmen, nämlich

$$\psi_{\text{I}} = \psi_{\text{II}}, \text{ Stetigkeit der tangentialen } \mathfrak{S} \quad (10)$$

$$\mathfrak{B}_{u\text{I}} = \mathfrak{B}_{u\text{II}}, \text{ Stetigkeit der normalen } \mathfrak{B}. \quad (10a)$$

Sie liefern, da sie für beliebige Werte $0 < v < \pi$ und $0 < \varphi < 2\pi$ gelten sollen

$$\begin{aligned} A - C f(u_0)/\mathfrak{Cof} u_0 &= -H \gamma c, \\ B - D g(u_0)/\mathfrak{Sin} u_0 &= -H \alpha c, \end{aligned} \quad (11)$$

$$A - \frac{\mu}{\mu_0} C f'(u_0)/\mathfrak{Sin} u_0 = \mathfrak{M} c - \frac{\mu}{\mu_0} H \gamma c, \quad (11a)$$

$$B - \frac{\mu}{\mu_0} D g'(u_0)/\mathfrak{Cof} u_0 = -\frac{\mu}{\mu_0} H \alpha c.$$

Zu dem Gleichungspaar (11a) bemerken wir, daß es nach (10a) und (2) entstanden ist aus

$$\mu_0 (-\partial \psi_{\text{I}} / \partial s_u + \mathfrak{M}_u) = -\mu \partial \psi_{\text{II}} / \partial s_u. \quad (11b)$$

Hier ist $\partial \psi / \partial s_u$ der Gradient von ψ in der u -Richtung, also gleich $\frac{\partial \psi}{\partial u} \frac{ds_u}{du}$; \mathfrak{M}_u ist die u -Komponente des Vektors $\mathfrak{M} = \mathfrak{M}_z$. Indem wir auf die Schemata

⁵⁾ Wir hätten natürlich den Term ψ_0 auch auf der rechten Seite der ersten Gl. (9) hinzufügen können. Das würde aber nur eine formale Abänderung der Konstanten A, B bedeuten, die für das Folgende nicht bequem wäre.

A und B in § 2 vorgehen, schreiben wir

$$\mathfrak{M}_u = c \mathfrak{M} \sin u_0 \cos v \frac{ds_u}{du}. \quad (11c)$$

Beim Einsetzen in (11b) bestätigt man, daß \mathfrak{M} nur in der ersten Gl. (11a) auftritt, welche dem (herausgehobenen) Faktor $\sin u_0 \cos v$ entspricht, nicht aber in der zweiten Gl. (11a), welche dem (ebenfalls herausgehobenen) Faktor $\cos u_0 \sin v \cos \varphi$ entspricht.

Ferner bemerken wir: Die beiden ersten Gln. (11) und (11a) verkoppeln A und C miteinander, die beiden zweiten B und D , so daß wir es bei der Ausrechnung nicht mit vierreihigen Determinanten zu tun haben, sondern nur mit zweireihigen. Das Ergebnis vereinfacht sich noch, wenn man von den aus den Diff. Gln. (7c) folgenden Beziehungen Gebrauch macht

$$\cos u f' - \sin u f = 1/\sin u, \quad \sin u g' - \cos u g = -2/\sin u. \quad (12)$$

Indem wir die Werte von $\cos u$ und $\sin u$ aus (5a) einsetzen, erhalten wir

$$\begin{aligned} A_1 A &= -H \gamma c - \frac{\mu_0}{\mu} \mathfrak{M} f(u_0) \frac{b^2}{c} \\ A_1 C &= \left(\frac{\mu - \mu_0}{\mu} H \gamma - \frac{\mu_0}{\mu} \mathfrak{M} \right) \frac{a b^2}{c^2} \end{aligned} \quad (13)$$

$$A_1 = 1 + \frac{\mu - \mu_0}{\mu} f(u_0) \frac{b^2}{c^2}$$

$$A_2 B = -2 H \alpha c$$

$$A_2 D = -\frac{\mu - \mu_0}{\mu} H \alpha \frac{a b^2}{c^2} \quad (14)$$

$$A_2 = 2 - \frac{\mu - \mu_0}{\mu} g(u_0) \frac{a b}{c^2}.$$

Nachdem die A, B, C, D bekannt sind, sind auch die Potentialfelder ψ_I, ψ_{II} durch (9) als stetige Ortsfunktionen bestimmt. Nur auf der Grenze zwischen I und II verhalten sich die normalen Gradienten von ψ_I und ψ_{II} unstetig, entsprechend dem hier induzierten Oberflächenmagnetismus.

§ 2. Vorbereitendes zur Berechnung des Drehmomentes

Das Oberflächenelement $d\sigma$ unseres Ellipsoids $u = u_0$ schreibt sich nach (6)

$$d\sigma = ds_u ds_\varphi = c^2 \sqrt{\sin^2 u_0 - \cos^2 v} \sin u_0 \sin v dv d\varphi \quad (15)$$

mit der Abkürzung

$$\sqrt{\sin^2 u_0 - \cos^2 v} = \sqrt{\cos^2 \varphi^2 u_0 - \cos^2 v}. \quad (15a)$$

Die äußere Normale auf $d\sigma$ hat die Richtung des Koordinatenzuwuchses du . Der Einheitsvektor in dieser Richtung heiße n . Der „modifizierte Spannungstensor“, Gl. (13) der vorangehenden Note, nimmt, vektoriell geschrieben, auf der äußeren Hüllfläche des Ellipsoids die Form an

$$\mathfrak{T} = \frac{\mu^2}{\mu_0} (\mathfrak{S} (\mathfrak{S} n) - \frac{1}{2} n \mathfrak{S}^2). \quad (16)$$

Das Drehmoment um einen äquatorialen Durchmesser (wir werden als solchen die y -Achse wählen) ist

$$N = \int [\mathbf{r} \times \mathbf{g}] d\sigma = \frac{\mu^2}{\mu_0} \int \{ [\mathbf{r} \times \mathbf{g}] (\mathbf{g} \cdot \mathbf{n}) - \frac{1}{2} [\mathbf{r} \cdot \mathbf{n}] \mathbf{g}^2 \} d\sigma. \quad (17)$$

Wir schreiben uns das orthogonale Schema auf, welches die beiden rechtwinkligen Achsenkreuze der Koordinatenänderungen dx, dy, dz und der Linienelemente ds_u, ds_v, ds_φ verbindet. Ebenso wie diese verhalten sich die Komponenten der Vektoren \mathbf{g} und \mathbf{M} nach den beiden Achsenkreuzen, was unser Schema ebenfalls zum Ausdruck bringt.

Schema A			0	0	\mathbf{M}_z
			\mathbf{g}_x	\mathbf{g}_y	\mathbf{g}_z
			dx	dy	dz
\mathbf{M}_u	\mathbf{g}_u	ds_u	α_{11}	α_{12}	α_{13}
\mathbf{M}_v	\mathbf{g}_v	ds_v	α_{21}	α_{22}	α_{23}
\mathbf{M}_φ	\mathbf{g}_φ	ds_φ	α_{31}	α_{32}	α_{33}

Das Schema kann von oben nach unten sowie von links nach rechts gelesen werden. Um die Werte der α_{ik} zu finden, berechne man nach (5)

$dx = c \operatorname{Coj} u \sin v \cos \varphi du + c \operatorname{Sin} u \cos v \cos \varphi dv - c \operatorname{Sin} u \sin v \sin \varphi d\varphi$,
und schreibe dafür mit Rücksicht auf (6a) und mit Benutzung der Abkürzung (15a)

$$dx = \frac{\operatorname{Coj} u \sin v}{\sqrt{\quad}} \cos \varphi ds_u + \frac{\operatorname{Sin} u \cos v}{\sqrt{\quad}} \cos \varphi ds_v - \sin \varphi ds_\varphi.$$

Daraus entnimmt man die α_{k1} in der ersten Spalte des Schemas B; ebenso erhält man die α_{k2} und α_{k3} aus den entsprechenden Formeln für dy, dz . Dadurch entsteht

Schema B	α_{k1}	α_{k2}	α_{k3}
α_{1k}	$\frac{\operatorname{Coj} u \sin v}{\sqrt{\quad}} \cos \varphi$	$\frac{\operatorname{Coj} u \sin v}{\sqrt{\quad}} \sin \varphi$	$\frac{\operatorname{Sin} u \cos v}{\sqrt{\quad}}$
α_{2k}	$\frac{\operatorname{Sin} u \cos v}{\sqrt{\quad}} \cos \varphi$	$\frac{\operatorname{Sin} u \cos v}{\sqrt{\quad}} \sin \varphi$	$\frac{\operatorname{Coj} u \sin v}{\sqrt{\quad}}$
α_{3k}	$-\sin \varphi$	$\cos \varphi$	0

Wir berechnen von hier aus nach Schema A

$$[\mathbf{r} \times \mathbf{g}]_y = z \mathbf{g}_z - x \mathbf{g}_x = (z \alpha_{11} - x \alpha_{13}) \mathbf{g}_u + (z \alpha_{21} - x \alpha_{23}) \mathbf{g}_v + (z \alpha_{31} - x \alpha_{33}) \mathbf{g}_\varphi; \quad (18)$$

nach Schema B und Gl. (15) haben wir

$$(z \alpha_{11} - x \alpha_{13}) d\sigma = c^3 \operatorname{Sin} u_0 \sin^2 v \cos v \cos \varphi dv d\varphi \quad (19a)$$

$$(z \alpha_{21} - x \alpha_{23}) d\sigma = c^3 \operatorname{Sin}^2 u_0 \operatorname{Coj} u_0 \sin v \cos \varphi dv d\varphi \quad (19b)$$

$$(z \alpha_{31} - x \alpha_{33}) d\sigma = -c^3 \sqrt{\quad} \operatorname{Sin} u_0 \operatorname{Coj} u_0 \cos v \sin v \sin \varphi dv d\varphi. \quad (19c)$$

Wir setzen (18) in (17) ein und berücksichtigen zugleich, daß

$$(\mathbf{g} \cdot \mathbf{n}) = \mathbf{g}_u, \quad \mathbf{g}^2 = \mathbf{g}_u^2 + \mathbf{g}_v^2 + \mathbf{g}_\varphi^2, \quad [\mathbf{r} \cdot \mathbf{n}]_y = z \alpha_{11} - x \alpha_{13}$$

ist. Wir erhalten dadurch

$$2 N \mu_0 [\mu^2 = \int (\mathfrak{H}_u^2 - \mathfrak{H}_v^2 - \mathfrak{H}_\varphi^2) (z \alpha_{11} - x \alpha_{13}) d\sigma \\ + 2 \int \mathfrak{H}_u \mathfrak{H}_v (z \alpha_{21} - x \alpha_{23}) d\sigma + 2 \int \mathfrak{H}_u \mathfrak{H}_\varphi (z \alpha_{31} - x \alpha_{33}) d\sigma] \quad (17)$$

und schreiben für die rechte Seite abkürzend

$$S_1 + S_2 + S_3. \quad (20a)$$

§ 3. Ausführung der Integrationen

Wir berechnen \mathfrak{H}_φ und \mathfrak{H}_v bequemerweise aus dem Potential ψ_I , was wegen der Stetigkeit der tangentialen \mathfrak{H} -Komponenten auf der äußeren ebenso wie auf der inneren Hüllfläche des Ellipsoids erlaubt ist. Wir erhalten so aus (9) und (6a)

$$\mathfrak{H}_\varphi = -\frac{d\psi_I}{ds_\varphi} = -\frac{\partial\psi_I}{\partial\varphi} \Big/ c \mathfrak{S} \sin u_0 \sin v = \frac{B}{c} \sin \varphi, \quad (21)$$

$$\mathfrak{H}_v = \frac{d\psi_I}{ds_v} = -\frac{\partial\psi_I}{\partial v} \Big/ c \sqrt{\quad} = (A \mathfrak{C} \mathfrak{O} \mathfrak{J} u_0 \sin v - B \mathfrak{S} \sin u_0 \cos v \cos \varphi) \Big/ c \sqrt{\quad}. \quad (22)$$

Andererseits können wir wegen der Stetigkeit der normalen \mathfrak{B} -Komponente auf der äußeren Hüllfläche schreiben

$$\mathfrak{H}_u = \frac{\mu_0}{\mu} \left(-\frac{d\psi_I}{ds_u} + \mathfrak{M} \right) = -\frac{\mu_0}{\mu} \left(\frac{\partial\psi_I}{\partial u} - \alpha_{13} c \sqrt{\quad} \mathfrak{M} \right) \Big/ c \sqrt{\quad} \\ = -\frac{\mu_0}{\mu} \{ (A - c \mathfrak{M}) \mathfrak{S} \sin u_0 \cos v + B \mathfrak{C} \mathfrak{O} \mathfrak{J} u_0 \sin v \cos \varphi \} \Big/ c \sqrt{\quad}. \quad (23)$$

Die Integration nach φ ist von 0 bis 2π zu erstrecken und läßt sich in allen drei Gliedern S_1, S_2, S_3 von (20a) ausführen. Betrachten wir zuerst S_3 . Mit Rücksicht auf (19c) besteht der Integrand aus einem Term mit $\sin^2 \varphi$ und einem mit $\sin^2 \varphi \cos \varphi$. Der letztere verschwindet, der erstere liefert den Faktor π . Daher wird

$$S_3 = \frac{\mu_0}{\mu} 2\pi (A - c \mathfrak{M}) B c \mathfrak{S} \sin^2 u_0 \mathfrak{C} \mathfrak{O} \mathfrak{J} u_0 \int_0^\pi \cos^2 v \sin v dv. \quad (24)$$

Im Integranden von S_2 treten Terme auf mit

$$\cos \varphi, \quad \cos^2 \varphi, \quad \cos^3 \varphi,$$

von denen nur der mittlere den Faktor π liefert, während die anderen verschwinden. Man erhält so:

$$S_2 = \frac{\mu_0}{\mu} 2\pi \mathfrak{C} \mathfrak{O} \mathfrak{J} u_0 \mathfrak{S} \sin^2 u_0 \left\{ (A - c \mathfrak{M}) B c \mathfrak{S} \sin^2 u_0 \int_0^\pi \frac{\cos^2 v}{\mathfrak{C} \mathfrak{O} \mathfrak{J}^2 u_0 - \cos^2 v} \sin v dv \right. \\ \left. - A B c \mathfrak{C} \mathfrak{O} \mathfrak{J}^2 u_0 \int_0^\pi \frac{\sin^2 v}{\mathfrak{C} \mathfrak{O} \mathfrak{J}^2 u_0 - \cos^2 v} \sin v dv \right\}. \quad (25)$$

Was schließlich S_1 betrifft, so kann man hier das Glied \mathfrak{H}_φ^2 streichen, weil es nach (21) den Faktor $\sin^2 \varphi$ enthält und nach (19a) mit $\cos \varphi$ zu multiplizieren ist.

Ferner braucht man in $\mathfrak{H}_u^2 - \mathfrak{H}_v^2$ wieder nur die Glieder mit dem Faktor $\cos \varphi$ zu berücksichtigen. Sie lauten

$$2 \left\{ \frac{\mu_0^2}{\mu^2} (A - c \mathfrak{M}) B + A B \right\} \mathfrak{C} \varphi u_0 \mathfrak{S} \sin u_0 \frac{\cos v \sin v}{c^2 (\mathfrak{C} \varphi^2 u_0 - \cos^2 v)} \cos \varphi$$

und liefern nach Multiplikation mit (19a) und Ausführung der φ -Integration

$$S_1 = 2\pi c \mathfrak{C} \varphi u_0 \mathfrak{S} \sin^2 u_0 \left\{ \frac{\mu_0^2}{\mu^2} (A - c \mathfrak{M}) B + A B \right\} \int \frac{\cos^2 v \sin^2 v}{\mathfrak{C} \varphi^2 u_0 - \cos^2 v} \sin v dv. \quad (26)$$

§ 4. Der Grenzfall des Stabmagneten

Es sei

$$b \ll a, \quad \mathfrak{S} \sin u_0 \ll \mathfrak{C} \varphi u_0, \quad u_0 \rightarrow 0.$$

Der Nenner des Integrals in Gl. (26) wird dann $1 - \cos^2 v$ und kann gegen den Faktor $\sin^2 v$ im Zähler gehoben werden; das Integral wird daraufhin gleich $\pi/3$. Indem man noch $\mathfrak{C} \varphi u_0$ und $\mathfrak{S} \sin u_0$ aus (5a) einsetzt, und das Volumen des Ellipsoids mit

$$V = \frac{4\pi}{3} a b^2$$

bezeichnet, erhält man aus (26)

$$S_1 = \frac{V}{c^2} \left\{ \frac{\mu_0^2}{\mu^2} (A - c \mathfrak{M}) + A \right\} B. \quad (27)$$

In Gl. (25) wird das zweite Integral in der Grenze $b \rightarrow 0$ gleich 2, das erste Integral wächst zwar wie $\log b/a$ ins Unendliche, verschwindet aber bei der Multiplikation mit dem davorstehendem Faktor $\mathfrak{S} \sin^2 u_0$. Daher wird

$$S_2 = -3 \frac{V}{c^2} \frac{\mu_0}{\mu} A B. \quad (28)$$

Aus (24) folgt ohne weiteres

$$S_3 = \frac{V}{c^2} \frac{\mu_0}{\mu} (A - c \mathfrak{M}) B. \quad (29)$$

Als Summe von (27), (28), (29) ergibt sich

$$\frac{V}{c^2} \left\{ \left(\frac{\mu_0^2}{\mu^2} - 3 \frac{\mu_0}{\mu} + \frac{\mu_0}{\mu} + 1 \right) A - \left(\frac{\mu_0^2}{\mu^2} + \frac{\mu_0}{\mu} \right) c \mathfrak{M} \right\} B = \frac{V}{c^2} \left\{ \frac{(\mu - \mu_0)^2}{\mu^2} A - \frac{\mu_0}{\mu^2} (\mu + \mu_0) c \mathfrak{M} \right\} B$$

und nach (20), (20a)

$$N = \frac{V}{c^2} \left\{ (\mu - \mu_0)^2 A - \mu_0 (\mu + \mu_0) c \mathfrak{M} \right\} \frac{B}{2 \mu_0}. \quad (30)$$

Die Werte von A und B entnimmt man aus (13) und (14). In der Grenze für $b \rightarrow 0$ hat man

$$A_1 = 1, \quad A_2 = 2 - \frac{\mu - \mu_0}{\mu} = \frac{\mu + \mu_0}{\mu}, \quad A = -H \gamma c, \quad B = -\frac{\mu}{\mu + \mu_0} 2 H \alpha c.$$

Aus (30) entsteht so:

$$N = V \mu \left\{ \mathfrak{M} + \frac{(\mu - \mu_0)^2}{\mu_0 (\mu + \mu_0)} H \gamma \right\} H \alpha. \quad (31)$$

Dies gilt für unseren modifizierten Spannungstensor. Hätten wir nicht mit diesem, sondern mit dem Maxwell-Tensor gerechnet, so wäre der Faktor $V \mu$ in (31)

durch $V\mu_0$ zu ersetzen. Dieser Unterschied eröffnet, wie schon in der vorangehenden Note betont, die Aussicht zwischen den beiden Tensoren und daher auch zwischen den zugehörigen Elementargesetzen [Gl. (11) und Gl. (4) der vorangehenden Note] experimentell zu unterscheiden.

Wir wollen ausdrücklich hervorheben, daß H in (31) das äußere Feld im Unendlichen bedeutet, nicht das Feld in der Umgebung des Stabes, das ja von jenem durchaus verschieden ist. Das Glied mit der Longitudinal-Komponente $H\gamma$ rührt offenbar von dem auf der Oberfläche des Ellipsoids influenzierten Magnetismus her und verstärkt das eingeprägte Moment \mathfrak{M} , wenn auch in praktisch unmerklichem Grade. Den bemerkenswerten Umstand, daß dieses Zusatzmoment von dem Quadrat der Suszeptibilität $\kappa = \mu - \mu_0$ abhängt und daher keine Unterscheidung zwischen Dia- und Paramagnetismus gestattet, hat schon Emil Cohn⁶⁾ betont. Die experimentelle Möglichkeit dieser Unterscheidung (Einstellung parallel oder senkrecht zum Felde) rührt nach Cohn vielmehr von Inhomogenitäten des Feldes her.

Es wäre unserer Meinung nach verkehrt, aus dem Auftreten des Faktors $\mu H\alpha = \mu H_n = B_n$ in (31) auf den dynamischen Charakter von B zu schließen, da B_n ja gar nicht die am Orte des Magneten wirkende Drehkraft, sondern, wie oben bei H hervorgehoben, das Feld im Unendlichen betrifft. Ebenso verkehrt wäre es, bei Zugrundelegung des Maxwell-Tensors; wo wie bemerkt in (31) μ_0 an die Stelle von μ tritt, aus dem dann auftretenden Faktor $\mu_0 H\alpha = \mu_0 H_n$ auf H als Drehkraft zu schließen. Das Drehmoment ist ja die Resultante aus Einzelwirkungen, die an den Oberflächen- und Raumelementen des Magneten angreifen; es ist viel zu kompliziert aufgebaut, um einen prinzipiellen Schluß auf den dynamischen Charakter von \mathfrak{B} oder \mathfrak{H} zuzulassen.

§ 5. Das Rotationsellipsoid beliebiger Form, insbesondere im Falle $\mu = \mu_0$

Um ein ähnlich einfaches Resultat wie für den Stabmagneten zu gewinnen, beschränken wir uns auf den Spezialfall $\mu = \mu_0$. Wir fassen die Glieder mit $(A - c\mathfrak{M})B$ in (26), (25) und (24) zusammen und erhalten mit der vorigen Abkürzung V

$$\frac{3}{2} \frac{V}{c^2} (A - c\mathfrak{M}) B \left\{ \int \frac{\cos^2 v \sin^2 v + \mathfrak{E} \sin^2 u_0 \cos^2 v}{\mathfrak{G} \sin^2 u_0 - \cos^2 v} \sin v dv + \int \cos^2 v \sin v dv \right\}. \quad (32)$$

Hier ist das erste Integral gleich dem zweiten wegen $\sin^2 v + \mathfrak{E} \sin^2 u_0 = \mathfrak{G} \sin^2 u_0 - \cos^2 v$. (32) vereinfacht sich daher zu

$$3 \frac{V}{c^2} (A - c\mathfrak{M}) B \int \cos^2 v \sin v dv = 2 \frac{V}{c^2} (A - c\mathfrak{M}) B. \quad (32a)$$

Andererseits liefern die beiden noch nicht berücksichtigten, mit dem Faktor AB behafteten Glieder in (25) und (26)

$$\frac{3}{2} \frac{V}{c^2} A B \left\{ \int \frac{\cos^2 v \sin^2 v}{\mathfrak{G} \sin^2 u_0 - \cos^2 v} \sin v dv - \int \frac{\mathfrak{G} \sin^2 u_0 \sin^2 v}{\mathfrak{G} \sin^2 u_0 - \cos^2 v} \sin v dv \right\}. \quad (33)$$

Dies zieht sich zusammen zu

$$-\frac{3}{2} \frac{V}{c^2} A B \int \sin^3 v dv = -2 \frac{V}{c^2} A B. \quad (33a)$$

⁶⁾ Das elektromagnetische Feld, 1. Aufl. Leipzig 1900 Kap. III im Anschluß an Gl. (43), wo aber nur eine allgemeine Abschätzung, keine eigentliche Berechnung des Drehmomentes gegeben wird.

Die Summe von (32a) und (33a) gibt

$$S_1 + S_2 + S_3 = -2 \frac{V}{c} \mathfrak{M} B. \quad (34)$$

Entnimmt man jetzt B aus (14), so ergibt sich

$$A_2 = 2, \quad B = -H \alpha c.$$

Aus (34) und (20), (20a) folgt daher mit $\mu = \mu_0$

$$N = \mu_0 V \mathfrak{M} H \alpha = \mu_0 V \mathfrak{M} H_n. \quad (35)$$

Dasselbe Resultat ergibt sich auch aus dem Maxwell'schen Tensor, weil dieser ja für $\mu = \mu_0$ mit unserem modifizierten Tensor (auch bei beliebiger Gestalt des Rotationsellipsoides) identisch ist.

Wenn $\mu \neq \mu_0$ ist, lassen sich die Integrale in (24), (25), (26) nicht so einfach zusammenziehen, wie es in Gl. (32) und (33) der Fall war, können aber einzeln elementar ausgerechnet werden, wobei Zusatzglieder von der Form

$$\sin^2 u_0 \log \frac{\cos u_0 + 1}{\sin u_0} = \frac{b^2}{c^2} \log \frac{a+c}{b}$$

entstehen. Diese kommen möglicherweise bei nicht sehr kleinem b als Korrekturen in Betracht und können vielleicht dazu dienen, ältere Betrachtungen über die scheinbare Lage der Pole in Magnetstäben zu erklären.

München, Institut für theoretische Physik.

(Bei der Redaktion eingegangen am 20. Mai 1950.)

Zur Definition der magnetischen Größen

Von Johannes Fischer

Inhaltsübersicht

Beitrag zu der Frage, ob die magnetischen Kraftwirkungen auf Magnetismus-träger durch $\mathfrak{K} = p_H \mathfrak{H}$, $\mathfrak{M} = [m_H \mathfrak{H}]$ oder durch $\mathfrak{K} = p_B \mathfrak{B}$, $\mathfrak{M} = [m_B \mathfrak{B}]$ zu beschreiben sind. — Die Gleichartigkeit hervorgebrachter und erlittener magnetischer Wirkungen von Leitungsströmen und von Magneten muß nicht notwendig dadurch zum Ausdruck kommen, daß beide Kraftwirkungen durch denselben magnetischen Kraftvektor, etwa \mathfrak{B} , dargestellt werden; die Gleichartigkeit kommt schon dadurch zum Ausdruck, daß die beiden Kraftwirkungen magnetisch ununterscheidbar ineinander überführt werden können, so daß man im Grunde genommen mit einem der beiden Kraftausdrücke auskommt. — Die Polstärke p_H ist grundsätzlich stets meßbar durch einen eindeutig definierten Vektorfluß außerhalb des Magneten, sie ist innere Eigenschaft des Magneten, unabhängig von der Permeabilität des umgebenden Mediums. Die Polstärke p_B ist von dieser abhängig, ihre genaue Bestimmung erfordert die nicht immer genau gegebene Kenntnis des magnetischen Feldes im Innern des Magneten. — Wenn man die elektrische Ladung als unabhängige Grundgröße ansieht, für die ein Erhaltungssatz gilt, so sind die elektrische Ladung q , das magnetische Moment m_B und die magnetische Polstärke p_B elektrische Quantitätsgrößen, m_H und p_H sind elektrische Intensitätsgrößen. Sieht man jedoch m oder p als unabhängige Grundgröße an, für die ein Erhaltungssatz gilt, so können m_H und p_H als magnetische Quantitätsgrößen aufgefaßt werden, aber die elektrische Ladung q ist dann magnetische Intensitätsgröße. Sind elektrische Ladung, magnetisches Moment und magnetische Polstärke Größen derselben Art (entweder Quantitäts- oder Intensitätsgrößen), dann gilt nur entweder für q , oder für m und p ein Erhaltungssatz. Gelten für die elektrische Ladung und ebenso für das magnetische Moment und die magnetische Polstärke Erhaltungssätze, so sind q , m und p nicht Größen derselben Art. Man muß eine der beiden möglichen Annahmen machen, um zu einer Entscheidung in der eingangs gestellten Frage kommen zu können.

1. Problemstellung

Der historische Weg zur quantitativen Beschreibung magnetischer Felder beginnt bei den Kraftwirkungen zwischen starren Körpern, die dauernd oder vorübergehend magnetisch polarisiert sind. Magnetische Mengen (Polstärken, Momente) werden vorausgesetzt, durch die Kraftwirkung wird die Stärke \mathfrak{H} des magnetischen Kraftfeldes definiert. Die Maxwell'schen Gleichungen verknüpfen das elektrische und das magnetische Feld als zwei unabhängig voneinander ge-

fundene und bestimmte Erscheinungskomplexe. Anders haben J. Wallot¹⁾ und ebenso E. Cohn²⁾ als ersten magnetischen Feldvektor die Induktion \mathfrak{B} definiert aus der Kraftwirkung auf das lineare Leitungsstrom-Element³⁾; indem man dann die magnetische Erregung \mathfrak{H} als Folgegröße aus dem Hauptvektor \mathfrak{B} durch $\mathfrak{H} = \mathfrak{B}/\mu\mu_0$ bestimmt, wird die Aussage des Durchflutungsgesetzes besonders einfach⁴⁾. Die Maxwellschen Gleichungen verknüpfen hier nicht die unabhängig voneinander definierten magnetischen und elektrischen Feldvektoren, sondern sie schließen die magnetischen Größen definierend an die gegebenen elektrischen an. Darum sind hier die magnetische Polstärke p und das magnetische Moment m nicht vorausgesetzte, sondern gefolgerte Größen (Hilfsgrößen). Die Analogie von Dauermagnet und stromdurchflossener Spule hat Wallot veranlaßt⁵⁾, eine solche Definition von p und m zu empfehlen, daß die Kraftwirkungen durch den Feldvektor \mathfrak{B} beschrieben werden, daß also die Kraft \mathfrak{K} und das Drehmoment \mathfrak{M} auf einen magnetisierten Körper bestimmt werden durch

$$\mathfrak{K} = p_B \mathfrak{B}', \quad \mathfrak{M} = [m_B \mathfrak{B}'], \quad (1,1)$$

während auf dem historischen Weg der Begriffsbildung definitionsweise gilt

$$\mathfrak{K} = p_H \mathfrak{H}, \quad \mathfrak{M} = [m_H \mathfrak{H}]. \quad (1,2)$$

A. Sommerfeld hat die Größen p_B , m_B befürwortet⁶⁾ und benutzt⁷⁾, vergleichen E. Hallén⁸⁾. H. Diesselhorst hat für die Untersuchung der Frage, welche der beiden Definitionen (1,1; 1,2) vorzuziehen sei, ein experimentelles Verfahren vorgeschlagen⁹⁾ und schließlich die Feldwirkungen in permeablen Medien untersucht¹⁰⁾. Dabei wurde aus den Faraday-Maxwellschen Spannungen das Drehmoment berechnet, das auf ein gleichmäßig magnetisiertes Rotationsellipsoid in einem homogenen Feld entsteht; eine Entweder-Oder-Entscheidung zwischen (1,1) und (1,2) wird durch diese Berechnung allein nicht erhalten. In einer Untersuchung, die hieran anknüpft, ist W. Döring¹¹⁾ zu dem Schluß gekommen, daß bei Wechsel des permeablen Mediums, in das ein Dauermagnet oder eine Stromschleife

¹⁾ J. Wallot, Elektrotechn. Z. **47**, 1009 (1926); **28**, 431, 1882 (1927).

²⁾ E. Cohn, Das elektromagnetische Feld, 2. Aufl. Berlin 1927, S. 71.

³⁾ Cohn hebt die Analogie dieser Definition mit der Definition der elektrischen Feldstärke \mathfrak{E} aus dem Coulombschen Gesetz hervor: dort $d\mathfrak{K} = dQ \cdot \mathfrak{E}$, hier $d\mathfrak{K} = [I dt \cdot \mathfrak{B}']$. Wallot benutzt im Gedankenexperiment zur Ausmessung eines Feldes \mathfrak{B}' die Kraftwirkung auf die kurze Seite einer rechteckigen, sehr langen starren linearen Stromschleife.

⁴⁾ Indessen kann \mathfrak{H} auch unabhängig, nicht als Folgegröße definiert werden. G. Mie definiert \mathfrak{H} ohne Voraussetzung über andere magnetische Größen durch einen Kompensationsversuch: Am untersuchten Feldort wird eine kleine enge Zylinderspule (Länge l , Einheitsvektor in axialer Richtung i dem Stromdichtevektor rechtswendig zugeordnet) nach Richtung i und Durchflutung zI so eingerichtet, daß ihr Inneres feldfrei wird; dann ist $\mathfrak{H} = -i \cdot zI/l$ definitionsweise und unabhängig von den Eigenschaften des magnetischen Nullindikators und deren etwaiger Änderung oder Beeinflussung (G. Mie, Lehrbuch der Elektrizität und des Magnetismus, 2. Aufl. Stuttgart 1941; auch die elektrische Verschiebungsdichte \mathfrak{D} kann unabhängig von \mathfrak{E} bestimmt werden: S. 104).

⁵⁾ a. a. O. S. 1014, 432, 1883.

⁶⁾ A. Sommerfeld, Z. techn. Physik **16**, 420 (1935); Ann. Physik **36**, 336 (1935).

⁷⁾ A. Sommerfeld, Vorlesungen über Theoretische Physik, Band III, Elektrodynamik, Wiesbaden 1948, S. 11, 88.

⁸⁾ E. Hallén, Transact. Roy. Inst. Techn. Stockholm Nr. 6 (1937).

⁹⁾ H. Diesselhorst, Elektrotechn. Z. **62**, 497 (1941).

¹⁰⁾ H. Diesselhorst, Ann. Physik (6) **3**, 11 (1948).

¹¹⁾ W. Döring, Ann. Physik (6) **6**, 69 (1949).

eingebettet ist, die Größe m_H immer dann ungeändert bleibt, „wenn auf Grund unserer physikalischen Einsicht keine Abhängigkeit von der Permeabilität des umgebenden Mediums zu erwarten ist“. Hiernach sollte man die Größe m_H als Definition des magnetischen Momentes bevorzugen.

Für die Bevorzugung von m_B und p_B sprechen ebenfalls gewichtige Gründe, wie besonders aus Sommerfelds Elektrodynamik (a. a. O.) hervorgeht. Wenn alle magnetischen Momente von umlaufenden Elektrizitätsmengen (Ringströmen, Ampèresche Äquivalenz) und sich drehenden Ladungen (Elektronenspin) herrühren, so erscheint es befriedigend, daß auch alle Kraftwirkungen durch denselben magnetischen Feldvektor \mathfrak{B} dargestellt werden. In dimensioneller Hinsicht wird m_B von durch das Produkt aus Fläche und Stromstärke, p_B durch das Produkt aus Geschwindigkeit und Ladung bestimmt. Daher sind m_B und p_B Quantitätsgrößen, ebenso wie die elektrische Ladung, während m_H und p_H Intensitätsgrößen sind. Es wird also zur Beurteilung der Definitionsgleichungen (1,1) und (1,2) beigetragen, wenn man die Darstellung der magnetischen Kraftwirkungen untersucht (in Abschnitt 2), ferner die Bestimmung der magnetischen Polstärke (in Abschnitt 3) und die Bestimmung der magnetischen Größen als Quantitäts- und als Intensitätsgrößen (in Abschnitt 4). Die Frage, ob (1,1) oder (1,2) vorzuziehen ist, hat offenbar nur Sinn, nachdem die zwei magnetischen Feldvektoren \mathfrak{B} und \mathfrak{H} als Größen anerkannt sind, die voneinander wesentlich verschieden sind.

2. Zur Darstellung der magnetischen Kraftwirkungen

Die Beziehungen

$$\operatorname{div} \mathfrak{B} = 0, \quad (2,1)$$

$$\operatorname{rot} \mathfrak{H} = \mathfrak{G}, \quad (2,2)$$

$$\mathfrak{B} = \mu \mu_0 \mathfrak{H} + \mathfrak{P} \quad (2,3)$$

bestimmen das magnetische Feld vollständig und eindeutig aus der Verteilung der μ , \mathfrak{P} , \mathfrak{G} . Hierbei ist insbesondere (2,3) Ausdruck der Erfahrung, nicht etwa Definitionsgleichung; \mathfrak{P} ist kennzeichnende (innere) Eigenschaft permanenter Magnete; bei Anwesenheit von Dauermagneten kann man nicht \mathfrak{B} einfach proportional zu \mathfrak{H} setzen, vielmehr ist \mathfrak{B} eine (fast) eindeutige lineare Funktion von \mathfrak{H} , nämlich (2,3)¹²⁾. — Eine vollständige Beschreibung des magnetischen Feldzustandes erfordert also zwei Feldvektoren. — Aus (2,1...3) ergeben sich Ausdrücke für die Kraft \mathfrak{f} auf die Volumeneinheit, die in Übereinstimmung stehen mit dem Energiehaltungssatz und mit den Faraday-Maxwellschen Spannungen des magnetischen Feldes der Form $\mathfrak{p} = \mathfrak{H} \cdot \mathfrak{B}' - \mathfrak{n} \cdot \mathfrak{H} \mathfrak{B}'/2$. Nimmt man an, daß μ , \mathfrak{P} , \mathfrak{G} mit der Materie starr verbundene Größen sind, so erhält man bekanntlich

$$\mathfrak{f} = [\mathfrak{G} \mathfrak{B}'] + \varrho_H \mathfrak{H} - \frac{\mu_0}{2} \mathfrak{H}^2 \nabla \mu = \mathfrak{f}_1 + \mathfrak{f}_2 + \mathfrak{f}_3, \quad (2,4)^{13)}$$

worin \mathfrak{B}' und ϱ_H definiert sind durch

$$\mathfrak{B}' = \mu \mu_0 \mathfrak{H} = \mathfrak{B} - \mathfrak{P}, \quad \varrho_H = \operatorname{div} \mathfrak{B}' = -\operatorname{div} \mathfrak{P}. \quad (2,4a)$$

¹²⁾ E. Cohn, a. a. O. S. 84; Physik. Z. **13**, 48 (1923).

¹³⁾ Man erhält $\mathfrak{f} = \mathfrak{f}_1 + \mathfrak{f}_2 + \mathfrak{f}_3$ mit $\mathfrak{f}_2 = (\mathfrak{B} \nabla) \mathfrak{H}$ und dazu eine Drehmomentdichte $\mathfrak{l} = [\mathfrak{P} \mathfrak{H}]$, wenn man starre elementare Ströme oder Magnete annimmt: H. Diesselhorst, Magnetische Felder und Kräfte, 1920, Neudruck Leipzig 1939, § 60, 88, 89.

Hiernach lassen sich nicht alle magnetischen Kraftwirkungen durch magnetische Feldvektoren \mathfrak{B}' oder \mathfrak{H} ausdrücken. \mathfrak{f}_3 ist zu dem Vektor $\nabla\mu$ proportional, nicht zu \mathfrak{B}' , nicht zu \mathfrak{H} ¹⁴⁾.

Der Wunsch, \mathfrak{f}_1 und \mathfrak{f}_2 , also die Kräfte auf Leitungsstromträger und die Kräfte auf Magnetismusträger, durch denselben Feldvektor ausgedrückt zu sehen, erscheint entbehrlich nach folgender Überlegung:

Wird eine Anordnung (\mathfrak{P} , $\mathfrak{G} = 0$) mit einer zweiten (\mathfrak{G} , $\mathfrak{P} = 0$) verglichen, und ist die Verteilung der \mathfrak{P} und \mathfrak{G} in beiden Fällen so, daß zutrifft

$$\mathfrak{G} = \text{rot} \frac{\mathfrak{P}}{\mu \mu_0}, \quad (2,5)$$

so kann bekanntlich nicht entschieden werden, ob das magnetische Feld von einem System von \mathfrak{G} oder von einem System von \mathfrak{P} hervorgebracht wird. In demselben äußeren Feld erfahren aber in diesem Fall auch die starren \mathfrak{G} -Träger dieselben Kräfte, wie die starren \mathfrak{P} -Träger, diese sind also allein in magnetischer Hinsicht in keiner Weise voneinander unterscheidbar. Ob ein magnetisches Feld von einem System starrer elementarer Ströme, oder von einem System starrer elementarer Permanentmagnete hervorgebracht wird, kann allein durch Untersuchung der ausgeübten und erlittenen magnetischen Kraftwirkungen nicht entschieden werden. Man kann ohne zusätzliche Erfahrung nicht angeben, ob die erste Ursache des magnetischen Feldes ein Wirbelsystem (\mathfrak{G}) oder ein Quellsystem (\mathfrak{P}) ist, man kann jedes von beiden ohne magnetische Wirkung und daher nicht unterscheidbar jeweils in das andere überführen. Hiernach kommt die ursächliche Gleichheit der magnetischen Kraftwirkungen nicht dadurch zum Ausdruck, daß Wirbelkräfte \mathfrak{f}_1 und Quellenkräfte \mathfrak{f}_2 durch denselben magnetischen Feldvektor dargestellt werden, sondern dadurch, daß Strömungsverteilung und Magnetverteilung in der ausgeübten und in der erlittenen Wirkung einander vollkommen gleichwertig sind, daß sie in magnetischer Hinsicht nicht voneinander unterschieden werden können. Anders gesagt: es erscheint hiernach deswegen nicht notwendig, daß die Kräfte auf Ströme (\mathfrak{f}_1) und die auf Magnete (\mathfrak{f}_2) durch denselben magnetischen Feldvektor ausgedrückt werden, weil im Grunde genommen nur einer der beiden Kraftausdrücke notwendig ist; dabei ist es vollkommen gleichgültig, welchen der beiden man als ersten wählt (diese Wahl bedeutet nur eine Entscheidung darüber, ob man die Beschreibung des Feldes mit der Bestimmung von \mathfrak{H} oder mit der von \mathfrak{B} beginnen will¹⁵⁾). Die Anerkennung von \mathfrak{B}' als magnetischem Hauptvektor (vgl. Abschnitt 1) führt also nicht zu der Notwendigkeit, die Kraftwirkungen auf dauermagnetische Körper gleichfalls durch \mathfrak{B}' darzustellen.

Für einen Ausdruck der Kraftdichte (und Drehmomentdichte), in welchem die Kraftwirkungen auf die entsprechend definierten Quellen des magnetischen Feldes proportional zu \mathfrak{B}' sind, nicht zu \mathfrak{H} , muß die Übereinstimmung mit der Erfahrung, mit dem Energieerhaltungssatz und mit den Faraday-Maxwellischen Spannungen bewiesen werden. Für den Ausdruck (2,4) darf dieser Beweis als erbracht gelten. Man kann noch allgemeiner fragen, ob für die Faraday-Maxwellischen Spannungen ein

¹⁴⁾ Im Fall $\mu = f(H)$ wird $\mathfrak{f}_3 = -\int_0^H \nabla\mu \cdot H \cdot dH$, also eine Funktion des Vektors $\nabla\mu$. — E. Cohn, a. a. O.

¹⁵⁾ Entweder erst \mathfrak{H} durch $\mathfrak{g}_H = \text{div} \mu \mu_0 \mathfrak{H}$, $\text{rot} \mathfrak{H} = 0$, hierauf $\mathfrak{B} = \mu \mu_0 \mathfrak{H} + \mathfrak{P}$, oder erst \mathfrak{B} durch $\text{rot} \mathfrak{B}/\mu \mu_0 = \mathfrak{G}$, $\text{div} \mathfrak{B} = 0$, hierauf $\mathfrak{H} = (\mathfrak{B} - \mathfrak{P})/\mu \mu_0$.

anderer Ausdruck in Betracht zu ziehen ist, als die bekannte, vor (2,4) angegebene Form. Dieser neue Ausdruck müßte dann ebenso, wie es jener alte tut, im Einklang stehen mit dem Energieerhaltungssatz und mit den aus der Erfahrung prüfbaren Folgerungen.

3. Zur Bestimmung der magnetischen Polstärke

Ein Körper heißt magnetisch polarisiert, wenn er entgegengesetzt gleiche magnetische Eigenschaften an verschiedenen Körperseiten zeigt. Deshalb gibt es auf der Oberfläche jedes Dauermagneten eine neutrale Flächenzone, auf der die normal zur Oberfläche gerichtete Feldkomponente sehr klein ist. Die neutrale Zone, auf deren Flächenelementen die normal gerichtete Feldkomponente verschwindet, wird im allgemeinen unendlich schmal sein. Wir nennen sie im Grenzfall kürzerhalb neutrale Linie L . Sie ist eine eindeutige, geschlossene Kurve auf der Körperoberfläche, und sie trennt diese in zwei Teilflächen S, S' ; sie ist die Kontur einer Querschnittsfläche σ , die das Körpervolumen in zwei Teilvolumina teilt. Die Teilflächen S, S' sind stets eindeutig, die Querschnittsfläche σ im Körperinneren bedarf näherer Festlegung. — Dann ist definitionsweise

$$p_H = \int_S B'_n df = \int_S \mu \mu_0 H_n df = \mu \mu_0 \int_S H_n df; \quad (3,1)$$

die Polstärke p_H ist der eine Teilfläche durchsetzende Fluß der magnetischen Induktion; μ ist die Permeabilität des umgebenden Mediums. Die Polstärke p_H kann somit grundsätzlich auch bei nicht einfachen Körperformen und Magnetisierungen durch Messung des magnetischen Flusses außerhalb vom Magneten bestimmt werden¹⁶⁾. Kenntnisse oder Annahmen über das magnetische Feld im Innern des Körpers sind nicht erforderlich.

Die Polstärke p_H ist von der Permeabilität μ des Mediums unabhängig, das den Magneten umgibt, und darf als innere Eigenschaft eines gegebenen permanenten Magneten betrachtet werden: Die Zustandsgleichung permanenter Magnete

$$\mathfrak{B} = \mu \mu_0 \mathfrak{H} + \mathfrak{P} \equiv \mu \mu_0 (\mathfrak{H} + \mathfrak{H}^*) \quad (3,2)$$

ist weitgehend experimentell belegt¹⁷⁾; μ und \mathfrak{P} sind konstante Größen, die erste kennzeichnet die Substanz, die zweite den magnetischen Zustand des Dauermagneten. Daher ist wegen $\text{div } \mathfrak{B} = 0$ die magnetische Flächendichte $\sigma = -\text{Div } \mathfrak{P} = -P_n$ und die Polstärke

$$p_H = \int_S P_n df \quad (3,3)$$

ist einzig bestimmt durch die innere Konstante \mathfrak{P} .

Die Polstärke p_B ist bestimmt als der Hüllenfluß

$$p_B = \oint H_n df; \quad (3,4)^{18)}$$

¹⁶⁾ R. W. Pohl, Elektrizitätslehre, 13. u. 14. Aufl. 1949, S. 99. — Wegen $\text{div } \mathfrak{B} = 0$ kommen für die Integration außer S und S' noch andere Flächen in Betracht, die L zur Kontur haben.

¹⁷⁾ R. Gans, Ann. Physik **22**, 481 (1907); **23**, 391 (1907); **27**, 1 u. 29 (1908); **29**, 301 (1909); **33**, 1065 u. a. (1910) — H. Laub, Arch. Elektrot. **16**, 481 (1926); W. Breitling, Arch. Elektrot. **35**, 1 (1941) u. Elektr. u. Masch.-Bau **61**, 315 (1943). — Zusammenfassend dargestellt in dem „Abriß der Dauermagnetkunde“ des Verfassers, Berlin, Göttingen, Heidelberg 1949.

¹⁸⁾ A. Sommerfeld, a. a. O.

die geschlossene Hüllfläche enthält den Körperquerschnitt σ und schließt sich beliebig um ein Teilvolumen, zum Beispiel über eine Teilfläche S, S' . Der Vergleich mit (3,1) ergibt den Zusammenhang¹⁹⁾

$$p_B - \int_{\sigma} H_n df = \frac{p_H}{\mu \mu_0} \quad (3,5)$$

Das bedeutet: die Polstärke p_B (und daher das Moment m_B) ist von der Permeabilität des umgebenden Mediums abhängig, jedoch nicht in einfacher Weise; bei nicht einheitlichen Eigenschaften der Umgebung des Magneten ist der Zusammenhang weniger einfach als (3,5). $\mu \mu_0$ ist nicht einfacher Proportionalitätsfaktor zwischen Polstärke p_B und Polstärke p_H , wie in früheren Diskussionen mehrfach angenommen wurde¹⁹⁾. Ferner: p_B enthält nach Definition (3,4) als einen Teil den Fluß von \mathfrak{H} durch den Körperquerschnitt σ , dessen Kontur die neutrale Linie L ist. Daher ist p_B nur dann genau bestimmt, wenn dieser Teilfluß angegeben werden kann, und angenähert, wenn er vernachlässigt werden kann gegen den Teilfluß $p_H/\mu \mu_0$ durch S oder S' .

Diese Vernachlässigung ist, ohne nähere Festlegung der Querschnittsfläche σ , bei allen längsmagnetisierten Körpern um so eher erlaubt, je gestreckter sie sind (auch wenn sie zum Beispiel hufeisenähnlich zusammengebogen sind). Dann sind p_B und p_H nach (3,4) einander proportional und gleichwertig, eine Entscheidung ist praktisch nicht nötig. Im gegebenen Einzelfall muß die Größe des Fehlers geprüft werden, der durch die Vernachlässigung des Flusses durch σ gemacht wird. Bei der homogen magnetisierten Kugel zum Beispiel berandet die neutrale Linie L als Äquator die Kreisfläche σ , die vom inneren Feld \mathfrak{H} normal durchsetzt wird, und der Teilfluß von \mathfrak{H} durch dieses σ ist halb so groß wie der Teilfluß durch eine der Halbkugelflächen S oder S' . Hier sind also p_B und $p_H/\mu \mu_0$ durchaus verschieden voneinander.

Bei einer Anzahl einfacher Körperformen und Körpermagnetisierungen hat das magnetische Feld im (homogen ausgefüllten) Außenraum eine Symmetrieebene, auf deren Elementen das Feld normal gerichtet ist (Stäbe und Scheiben, die längs oder die quer zur Achse symmetrisch magnetisiert sind). Die neutrale Linie L ist dann die Kontur der Symmetrieebene auf der Körperoberfläche, und σ ist der ebene Querschnitt, dessen Rand L ist. Der Fluß durch σ kann z. B. durch Entwerfen eines Feldbildes angenähert bestimmt werden. Bei allen parallel zu einer Achse homogen magnetisierten Rotationsellipsoiden kann der Fluß von \mathfrak{H} durch σ berechnet werden, er ist proportional zum Entmagnetisierungsfaktor. In diesem Fall kann das magnetische Moment unmittelbar bestimmt werden.

4. Zur Bestimmung der elektrischen und magnetischen Größen als Quantitäts- und als Intensitätsgrößen

a) Nicht nur für die magnetische Menge sind Definitionen möglich, die Größen von verschiedener physikalischer Art ergeben, sondern auch für die elektrische Ladung. Das ist für die Einteilung der Größen in Quantitäts- und Intensitätsgrößen wichtig.

Wir stellen im elektrischen Feld den Vektor der meßbaren mechanischen Kraft dar durch ein Produkt aus zwei Faktoren; den einen schreiben wir dem Meßkörper zu, den anderen dem Feldpunkt. Wird bei Wechsel des (homogenen) Mediums, in welchem das elektrische Feld besteht, eine Veränderung der Kraft gemessen, so kann man allein aus dieser Wahrnehmung natürlich nicht schließen, ob man dem einen oder dem anderen Faktor des Produktes oder beiden eine Veränderung zuzuschreiben hat. Dazu bedarf es einer zusätzlichen Erfahrung oder Voraussetzung

¹⁹⁾ Im Anschluß an die a. a. O. genannten Ausführungen von J. Wallot.

(Definition, Forderung). Eine solche ist die Erhaltung der elektrischen Ladung: die Ladung des Meßkörpers und das elektrisch polarisierte umgebende Medium wirken nicht so zusammen, daß bei einem Wechsel des Mediums eine Änderung der Größe der Ladung des Probekörpers anzunehmen wäre. Man kann diese Eigenschaft, unabhängig davon, ob sie einer Voraussetzung (Annahme, Definition) entfließt, oder ob sie aus der Erfahrung folgt, kürzshalber den Erhaltungssatz nennen. Für die Masse, die elektrische Ladung q , die magnetische Polstärke p_H und das magnetische Moment m_H gilt ein Erhaltungssatz, aber nicht für p_B und m_B .

Die Kraft zwischen zwei ruhenden Trägern elektrischer Ladungen q_E, q'_E in großem Abstand r voneinander ändert sich von ihrem Vakuumwert

$$\mathfrak{K}_0 = \frac{q_E q'_E}{\epsilon_0 \cdot 4 \pi r^2} \cdot \frac{r}{r} \quad (4,1)$$

bei Ausfüllung des Feldraumes mit einem Isolator $\epsilon > 1$ in $\mathfrak{K} = \mathfrak{K}_0/\epsilon$, wenn an den Ladungsträgern und ihrem Abstand nichts geändert wird. Da wir in diesem Fall Grund haben, die Elektrizitätsmengen für unverändert zu halten, drücken wir die Beobachtung aus durch

$$\mathfrak{K}_0 = q_E \mathfrak{C}_0, \quad \mathfrak{K} = q_E \mathfrak{C}, \quad \mathfrak{C} = \mathfrak{C}_0/\epsilon, \quad q_E = \text{const}; \quad (4,2)$$

die Änderung des Produktes wird also nicht dem Faktor zugeschrieben, der dem Meßkörper eigentümlich ist, sondern dem, der dem Feldpunkt eigentümlich ist. Weniger zweckmäßig, aber gleichfalls möglich ist es, dieselbe Beobachtung darzustellen durch

$$\mathfrak{K}_0 = (q_D)_0 \mathfrak{D}, \quad \mathfrak{K} = q_D \mathfrak{D}, \quad q_D = (q_D)_0/\epsilon, \quad \mathfrak{D} = \text{const}; \quad (4,3)$$

hier wäre also der dem Feldpunkt eigentümliche Faktor unabhängig von ϵ angenommen, und es würde sein

$$q_D = \frac{q_E}{\epsilon \epsilon_0} \quad (4,4)$$

abhängig von der Dielektrizitätskonstante des umgebenden Mediums und selbst im Fall $\epsilon = 1$ eine Größe von anderer Art, als q_E . Die Eigenschaften der zwei Faktoren sind eben nicht durch die Beobachtungen \mathfrak{K}_0 und \mathfrak{K} allein bestimmt. Nicht nur beim magnetischen Feld, auch beim elektrischen, sind verschiedene Definitionen für den dem Meßkörper zugeschriebenen Faktor möglich, und damit verschiedene Definitionen für die Quellen des Feldes und für die Beschreibung der Kraftwirkung durch den einen oder den anderen Feldvektor. Die Entscheidung wird erst durch eine zusätzliche Erfahrung oder Annahme ermöglicht. Eine solche ist insbesondere der beschriebene Erhaltungssatz — entweder als Erfahrungstatsache, oder als Voraussetzung (Postulat, Definition).

b) Da zugunsten von p_B und m_B die Eigenschaft als Quantitätsgrößen angeführt wird, untersuchen wir, ob die Einteilung aller Größen in Quantitäts- und in Intensitätsgrößen eindeutig ist. Dazu beginnen wir mit einer Definition dieser Begriffe:

Die Einheiten aller elektrischen und magnetischen Größen können bekanntlich als Potenzenprodukte von höchstens vier voneinander unabhängigen Grundeinheiten ausgedrückt werden. Drei von diesen seien die Einheiten $[L]$, $[T]$, $[W]$ für Länge, Zeit, Energie als unabhängige Grundgrößen. Wir setzen zunächst als vierte unabhängige Grundgröße die elektrische Ladung Q , voraussetzend, daß ihre Größe vom Wechsel des isolierenden Mediums, das den Ladungsträger umgibt, nicht abhängt. Daher ist für die Ladungseinheit $[Q]$ die Vereinbarung eines Normalmediums nicht nötig. Dann sind nach üblicher Definition elektrische Quantitätsgrößen alle die, deren Einheiten den Faktor $[Q]$ enthalten, und elektrische Intensitätsgrößen alle die, deren Einheiten den Faktor $[W]/[Q]$ enthalten. Da durch $U = W/Q$ die elektrische Spannung definiert wird, kann man die elektrischen

Intensitätsgrößen auch elektrische Spannungsgrößen, die elektrischen Quantitätsgrößen auch elektrische Ladungsgrößen nennen. Mit Hilfe der Gleichungen der Elektrizitätslehre ergeben sich, indem man von den elektrischen Größen ausgehend zu den magnetischen fortschreitet, als

$$\left. \begin{array}{l} \text{el. Ladungsgrößen (elektrische Quantitätsgrößen)} \\ q_E \quad \mathfrak{D} \quad \mathfrak{G} \quad I^{(20)} \quad \mathfrak{S} \quad m_B, \\ \text{el. Spannungsgrößen (elektrische Intensitätsgrößen)} \\ U \quad \mathfrak{E} \quad \Phi^{(20)} \quad \mathfrak{H}; \end{array} \right\} \quad (4,5)$$

ϵ_0 und $1/\mu_0$ sind Quotienten: Ladungsgröße zu Spannungsgröße; die erste Maxwell'sche Gleichung verknüpft elektrische Ladungsgrößen miteinander, die zweite elektrische Spannungsgrößen. In dem Ausdruck für die elektrische Kraft auf ruhende Ladungsträger $\mathfrak{K} = q_E \mathfrak{E}$ ist der dem Meßkörper zugeschriebene Faktor elektrische Quantitätsgröße (Ladungsgröße), der dem Feldpunkt zugeschriebene Faktor elektrische Intensitätsgröße (Spannungsgröße). Fordert man dieselbe Zusammensetzung für die Ausdrücke der magnetischen Kraftwirkungen, so können diese nach (4,5) nur so gebaut sein: $\mathfrak{M} = [m_B \mathfrak{H}']$ und $\mathfrak{K} = p_B \mathfrak{H}'$, denn \mathfrak{H}' , nicht \mathfrak{H} , ist elektrische Intensitätsgröße (Spannungsgröße).

Zu den Größen m_B und p_B hat der im 1. Abschnitt genannte Weg geführt, auf dem, von den elektrischen Größen ausgehend, die Maxwell'schen Gleichungen zur Definition der magnetischen Größen dienen. Dem entspricht es, daß die elektrische Ladung als unveränderliche, unabhängige Grundgröße gesetzt wird, um alle elektrischen und magnetischen Größen in elektrische Quantitätsgrößen und elektrische Intensitätsgrößen einzuteilen. Dies ist das Vorgehen von G. Mie und A. Sommerfeld.

Man kann auch den umgekehrten Weg gehen, nämlich ausgehend von den magnetischen Größen zu den elektrischen kommen. Man wird zweifellos die Zweckmäßigkeit, aber kaum die formale Zulässigkeit dieses Vorgehens in Frage stellen können. Dem gezeigten Vorgehen entspricht es, daß man hier von dem magnetischen Moment oder der Polstärke p als unabhängiger Grundgröße ausgeht und für diese einen Erhaltungssatz voraussetzt. $[P][L]$ sei eine Einheit der unveränderlichen, unabhängigen Grundgröße: magnetisches Moment m , und durch $V = W/p$ werde die magnetische Spannung definiert. Dann bezeichnen wir als magnetische Polstärkegrößen alle die, deren Einheiten den Faktor $[P]$, und als magnetische Spannungsgrößen alle die, deren Einheiten den Faktor $[W]/[P]$ enthalten. Mit den Gleichungen der Elektrizitätslehre ergeben sich dann, indem man von den magnetischen Größen zu den elektrischen fortschreitet, als

$$\left. \begin{array}{l} \text{magnetische Polstärkegrößen} \\ m_H \quad \mathfrak{H} \quad \Phi \quad \mathfrak{E} \quad U, \\ \text{magnetische Spannungsgrößen} \\ \mathfrak{S} \quad I \quad \mathfrak{G} \quad \mathfrak{D} \quad q_E. \end{array} \right\} \quad (4,6)$$

²⁰⁾ I elektrische Stromstärke, Φ magnetischer Induktionsfluß. — Für die elektrische Stromstärke hatte man das Formelzeichen I an Stelle des früheren c vereinbart, um an die Bezeichnung „Intensität“ dieser Quantitätsgröße zu erinnern.

μ_0 und $1/\epsilon_0$ sind Quotienten: magnetische Polstärkegröße zu Spannungsgröße; die erste Maxwellsche Gleichung verknüpft magnetische Spannungsgrößen miteinander, die zweite Polstärkegrößen. In den Ausdrücken der magnetischen Kraftwirkungen $\mathfrak{M} = [m_H \mathfrak{H}]$ und $\mathfrak{K} = p_H \mathfrak{H}$ ist der dem Meßkörper zugeschriebene Faktor Polstärkegröße, der dem Feldpunkt zugeschriebene Spannungsgröße. Fordert man dieselbe Zusammensetzung auch für den Ausdruck der elektrischen Kraftwirkungen auf ruhende Ladungsträger, so kann dieser nach (4,6) nur so gebaut sein $\mathfrak{K} = q_D \mathfrak{D}$, denn \mathfrak{D} , nicht \mathfrak{E} , ist magnetische Spannungsgröße. Bezeichnet man die Polstärkegrößen als magnetische Quantitätsgrößen, die magnetischen Spannungsgrößen als magnetische Intensitätsgrößen, so sind die elektrischen Quantitätsgrößen in (4,5) magnetische Intensitätsgrößen in (4,6), und die elektrischen Intensitätsgrößen in (4,5) sind magnetische Quantitätsgrößen in (4,6). Erklärt man anders die Polstärkegrößen zu magnetischen Intensitätsgrößen, die magnetischen Spannungsgrößen zu magnetischen Quantitätsgrößen, so sind die elektrischen Quantitätsgrößen in (4,5) magnetische Quantitätsgrößen in (4,6) und die elektrischen Intensitätsgrößen in (4,5) sind magnetische Intensitätsgrößen in (4,6).

Man darf folgern: Die Einteilung in Quantitätsgrößen und Intensitätsgrößen ist nicht eindeutig, sondern abhängig vom wählbaren Ausgangspunkt. Man kann Quantitätsgrößen und Intensitätsgrößen elektrischer und solcher magnetischer Art bestimmen. G. Mie und A. Sommerfeld haben aus einleuchtenden Gründen elektrische Quantitäts- und Intensitätsgrößen allein in Betracht gezogen. Die Ausdrücke $\mathfrak{M} = [m_B \mathfrak{B}']$ und $\mathfrak{K} = q_D \mathfrak{D}$ erscheinen beide als Folgen des gewählten Ausgangspunktes, $[m_B \mathfrak{B}']$ bringt nicht ein besonderes Merkmal der magnetischen Erscheinungen zum Ausdruck.

Das erste Vorgehen entspricht Maxwells „elektrischem System“, das zweite seinem „magnetischen System“ der Einheiten²¹⁾. Die elektrische Ladung, die Grundgröße des ersten ist, wird dabei von Maxwell als Flächenintegral des Verschiebungsvektors \mathfrak{D} definiert²²⁾, die magnetische Polstärke, die Grundgröße des zweiten ist, als Flächenintegral des Induktionsvektors \mathfrak{B}' ²³⁾. Man darf darum mit großer Sicherheit annehmen, daß Maxwell nicht nur für die elektrische Ladung einen Erhaltungssatz vorausgesetzt hat, sondern auch für die magnetische Polstärke. Dann kann man auch dieses „magnetische System“ und seine Voraussetzung $\mathfrak{M} = [m_H \mathfrak{H}]$ nicht verwerfen²³⁾.

Ein Erhaltungssatz besteht definitionsweise für die elektrische Ladung q_E (und das elektrische Moment e_E), für das magnetische Moment m_H und die magnetische Polstärke p_H , nicht für m_B , p_B , q_D . Andererseits sind, wie man auch die Einteilung treffen mag — in elektrische Ladungs- und Spannungsgrößen, oder in magnetische Polstärke- und Spannungsgrößen — Größen derselben Art q_E , m_B , p_B einerseits, m_H , p_H , q_D andererseits. Hiernach besteht offenbar

²¹⁾ J. Clerk Maxwell, Treatise on electricity and magnetism, Oxford 1873, Art. 623.

²²⁾ a. a. O. Art. 622.

²³⁾ Es liegt nahe, ein gemischtes, symmetrisches System aufzustellen, für das dann fünf voneinander unabhängige Grundgrößen erforderlich sind. In ihm kann man die beiden Ausdrücke $\mathfrak{K} = q_E \mathfrak{E}$ und $\mathfrak{K} = p_H \mathfrak{H}$ unabhängig voneinander an den Anfang stellen. Dabei wird eine neue physikalische Größe erforderlich, die als Faktor in den beiden Maxwellschen Gleichungen auftritt. Sie spielt eine ähnliche Rolle wie der Ausgleichfaktor in den Zahlenwertgleichungen der symmetrischen (gemischten) CGS-Systeme. Der Verfasser verdankt den Hinweis auf dieses System und seine Eigenschaften anregendem Briefwechsel mit Herrn Prof. Dr. R. Fleischmann, Hamburg.

für die Begriffe elektrische Ladung q , wie auch magnetisches Moment m und magnetische Polstärke p die Alternative: Entweder fordert man, daß q , m , p Größen derselben Art sein müssen; dann gilt nur entweder für die elektrischen oder für die magnetischen Größen ein Erhaltungssatz. Oder man setzt Erhaltungssätze sowohl für q , als auch für m , p voraus; dann sind die elektrischen und die magnetischen Größen nicht Größen derselben Art²⁴⁾.

Damit ist die Frage nach der Definition des magnetischen Momentes auf die Entscheidung zwischen zwei voneinander unabhängigen Voraussetzungen zurückgeführt. Ohne daß eine dieser beiden möglichen Annahmen gemacht wird, erscheint es nicht möglich, zwischen den beiden möglichen Definitionen des magnetischen Momentes und der magnetischen Polstärke durch \mathfrak{B}' oder durch \mathfrak{B} nach (1, 2) oder (1, 1) zu entscheiden.

²⁴⁾ Dieser Schluß gilt unter der Voraussetzung, daß die Anzahl der unabhängigen Grundgrößen vier ist. Bei fünf unabhängigen Grundgrößen, darunter einer elektrischen und einer magnetischen, kann die Einteilung aller Größen auf die gezeigte einfache Weise nicht mehr eindeutig vollzogen werden; die Eindeutigkeit erfordert eine zusätzliche Aussage. Bei Annahme dreier unabhängiger Grundgrößen geht gleichfalls die Eindeutigkeit der Einteilung verloren.

Karlsruhe, Schirmerstr. 6.

(Bei der Redaktion eingegangen am 6. Juni 1950.)

Bolometer geringer Trägheit

Von M. Czerny, W. Kofink und W. Lippert

(Mit 6 Abbildungen)

Inhaltsübersicht

Bekanntlich wird das optische Verhalten der Metalle gegenüber hinreichend langwelliger ultraroter Strahlung durch die spezifische Leitfähigkeit der Metalle bestimmt. W. Woltersdorff zeigte theoretisch in seiner Dissertation (1934)¹⁾, daß es für jedes Metall eine gewisse Schichtdicke gibt, bei der eine ultrarote Strahlung zu 50% in der Schicht absorbiert wird, während je 25% reflektiert, beziehungsweise hindurchgelassen werden. Der Effekt ist bemerkenswerterweise unabhängig von der Wellenlänge, soweit es sich um die oben gekennzeichnete langwellige ultrarote Strahlung handelt. Die in Frage kommenden Schichtdicken der Metalle liegen in der Größenordnung von $0,01 \mu$ und weniger. Woltersdorff zeigte ferner experimentell, daß sich durch Aufdampfen im Vakuum Metallschichten herstellen lassen, die in mehr oder minder guter Annäherung diese theoretischen Voraussagen erfüllen.

Im Anschluß an dessen Arbeit konstruierte G. Heins in den Jahren 1936/37 in einer leider unvollendeten Dissertation einen Typ von Bolometer, bei dem eine Metallschicht der Woltersdorffschen Dicke auf einer sehr dünnen Zelluloid-Membran aufgedampft war. Da die dünne Metallschicht gleichzeitig als Absorber und als Bolometerschicht verwandt wurde, erzielte er ein Empfangsinstrument sehr geringer Trägheit, das er erfolgreich bei einer Wechsellicht-Methode im langwelligen Ultrarot einsetzen konnte.

Der Bau solcher Bolometer wurde dann 1940 im Physikalischen Institut der Universität Frankfurt a. M. wieder aufgenommen. Es wurden dabei einige systematische Untersuchungen angestellt im wesentlichen mit dem Ziel, ein Bolometer besonders geringer Trägheit zu entwickeln. Über diese mehrere Jahre zurückliegenden Arbeiten soll im Folgenden berichtet werden. Unabhängig von uns und ohne unser Wissen wurde bei der Firma Zeiss in Jena in gleicher Richtung gearbeitet. Wir verwandten im wesentlichen nur Bi. Soweit bekannt, wurde bei Zeiss mit Sb gearbeitet.

§ 1. Die Absorption der Strahlung im Bolometer

Das Absorptionsvermögen einer dünnen Metallschicht für hinreichend langwellige ultrarote Strahlung ist nach Woltersdorff unabhängig von der Wellenlänge und wird bestimmt durch den Ausdruck

$$A = \frac{4\pi\sigma d/c}{(1 + 2\pi\sigma d/c)^2}. \quad (1)$$

¹⁾ H. Murmann, Z. Physik **54**, 741 (1929); R. B. Barnes u. M. Czerny, Phys. Rev. **38**, 338 (1931); W. Woltersdorff, Z. Physik **91**, 230 (1934).

Hierin bedeutet d die Schichtdicke in cm, σ die spezifische Leitfähigkeit des Metalls in elektrostatischen Einheiten und c die Lichtgeschwindigkeit. Der Ausdruck führt im Maximum zu einem Absorptionsvermögen von 50%, wenn nämlich

$$d \cdot \sigma = c/2\pi \quad (2)$$

ist.

Wenn eine Metallschicht der Dicke d quadratische Oberfläche mit der Seitenlänge l hat und in Richtung einer Quadratseite von Strom gleichmäßig durchflossen wird, so hat sie den Widerstand

$$\frac{l}{ld} \frac{1}{\sigma} = \frac{1}{\sigma d}.$$

Bezeichnet man diesen Widerstand als ihren Quadrat-Widerstand, so folgt aus Obigem, daß eine Schicht aus einem beliebigen Metall dann 50% absorbiert, wenn sie den Quadratwiderstand $2\pi/c$ in elektrostatischen Einheiten hat. Das entspricht in techn. Einheiten dem Wert $R^* = 188,5$ Ohm. Dieser Wert spielt daher bei den folgenden Ausführungen eine wesentliche Rolle.

Man kann den in der Metallschicht zur Absorption kommenden Anteil noch erhöhen, wenn man hinter der Schicht einen Reflektor aufstellt. Der durch die Schicht zunächst hindurchgehende Teil der Strahlung wird dann in Mehrfach-Reflexion zwischen Reflektor und Schicht noch mehrmals der Schicht zur Absorption angeboten. Wenn man dabei von Interferenzeffekten absieht, so ergibt eine hier nicht näher ausgeführte Rechnung²⁾, daß im günstigsten Falle der in der

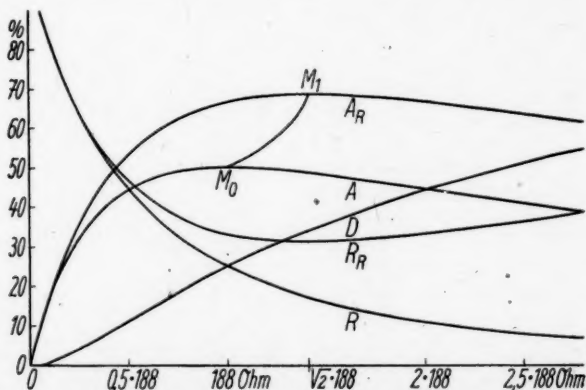


Abb. 1. Auf der Abszisse: Quadratwiderstand einer Metallschicht in Ohm. Auf der Ordinate: Prozente. Die Kurven A , D , R geben Absorption, Durchlässigkeit und Reflexionsvermögen einer Metallschicht für langwellige ultrarote Strahlung. Die Kurven A_R und R_R geben den absorbierten und reflektierten Anteil, wenn hinter der Metallschicht ein hundertprozentiger Reflektor steht. Der Kurvenzug von M_0 nach M_1 zeigt, wie das Maximum des absorbierten Anteils ansteigt, wenn hinter der Schicht Reflektoren stehen, deren Reflexions-Vermögen von 0 auf 100% ansteigt

²⁾ Wenn der hinter der Membran aufgestellte Reflektor 100% reflektiert, so wird der in der Membran absorbierte Anteil $A_R = \frac{4v}{(1+v)(2+v)}$ mit $v = \frac{R_{\square}}{R^*}$ und $R_{\square}^* = 188$ Ohm.

Schicht absorbierte Anteil von 50% auf 68,6% ansteigt. Dieser günstigste Fall tritt der Rechnung nach bei Schichten ein, die einen Quadratwiderstand von $\sqrt{2} \cdot 188 = 266 \text{ Ohm}$ aufweisen.

Die Zusammenhänge sind in Abb. 1 graphisch dargestellt. Die Kurven sind vor allem insofern von Bedeutung, als sie zeigen, daß der Quadratwiderstand einer Schicht in ziemlich weitem Maße variiert werden darf, ohne daß der Absorptionsanteil sehr unter seinen optimalen Wert absinkt.

Die hier zugrunde gelegten theoretischen Betrachtungen setzen voraus, daß die Strahlung senkrecht auf die Membran einfällt. Kürzlich ist eine Arbeit von G. Hettner³⁾ erschienen, in der die Rechnungen auch für schiefen Strahleneintritt verallgemeinert werden. Es sei hier nur auf diese Arbeit hingewiesen.

Die Metallschichten waren auf Zelluloid-Membranen von etwa $0,04 \mu$ Dicke aufgedampft. Diese Trägerschichten dürften höchstens dort etwas zur Gesamtabsorption des Bolometers beitragen, wo die Zelluloid-Schicht ausgeprägte Stellen selektiver Absorption besitzt.

§ 2. Über die Wahl des Metalls

Da für das Auftreten maximaler Absorption bei einem beliebigen Metall das Produkt aus Schichtdicke und spezif. Leitfähigkeit einen bestimmten Wert annehmen muß, so sollte man ein Metall mit hoher Leitfähigkeit wählen, um mit besonders geringer Schichtdicke auszukommen und auf diese Weise besonders kleine thermische Trägheit zu erzielen. Unsere Versuche, in dieser Hinsicht das günstigste Metall auszuwählen, sind wohl noch nicht als abschließend zu bezeichnen, aber zunächst wurden wir durch Vorversuche auf Bi, also gerade ein Metall mit geringer Leitfähigkeit, gedrängt. Der Grund dafür ist der, daß noch andere Eigenschaften des Metalls bei der Wahl beachtet werden müssen.

Rechnet man für Ag die Schichtdicke aus, bei der sich ein Quadratwiderstand von $188,5 \text{ Ohm}$ ergibt, so kommt man zu einem Wert von $0,85 \cdot 10^{-8} \text{ cm}$, wenn man den spez. Widerstand $\rho = 1,6 \cdot 10^{-6} \text{ Ohm cm}$ des massiven Metalls zugrunde legt. Bei Bi ergibt sich entsprechend eine Schichtdicke von $62 \cdot 10^{-8} \text{ cm}$, wenn man mit $\rho = 1,17 \cdot 10^{-4} \text{ Ohm cm}$ rechnet. Man ersieht aus diesen Zahlen, daß beim Ag die Schichtdicke so klein wird, daß auf keinen Fall mehr ein normales metallisches Verhalten zu erwarten ist. Selbst beim Bi fanden wir durch Wägung, daß Schichten, die einen Quadratwiderstand von 188 Ohm zeigten, etwa 6 bis 7 mal größere Masse besaßen, als sie nach obiger Rechnung zu erwarten ist. Da keinesfalls eine übernormal große Dichte bei so dünnen Schichten zu erwarten ist, ist also der spez. Widerstand unserer Bi-Schichten mindestens 6 bis 7mal größer als der des massiven Metalles.

Damit steht auch im Zusammenhang, daß die dünnen Bi-Schichten im Gegensatz zum massiven Bi einen negativen Temperatur-Koeffizienten des Widerstandes besaßen. Sein Absolutwert war etwa halb so groß wie der von massivem Bi.

Für die Verwendung von Bi sprach die nicht gar zu hohe Temperatur, bei der dieses Metall im Vakuum verdampft. Das ist wesentlich, um eine Zerstörung der zarten Zelluloid-Schichten zu verhindern. Die Haltbarkeit der Bi-Bolometer hat sich als gut erwiesen. Schließlich ist noch hervorzuheben, daß bei den Bolometern die Wärmekapazität der Bi-Schicht noch nicht größer war als die der $0,04 \mu$ dicken Trägerschicht, wie im Folgenden noch besprochen wird.

³⁾ G. Hettner, *Optik* 1, 1 (1946).

§ 3. Aufbau des Bolometers

Als Normaltyp wurde ein Doppel-Bolometer gebaut mit 2 möglichst gleichartigen Schichten. Die eine ist nach vorn der aufzunehmenden Strahlung, die zweite nach hinten der Rückwand des Bolometergehäuses zugekehrt. Jede Schicht besteht aus Bi, hat etwa 100 bis 300 Ohm Quadratwiderstand und ist durch Bedampfen einer Zelluloid-Membran im Hochvakuum hergestellt. Als Träger der Membran und gleichzeitig als Elektrode dient ein quadratisches Rähmchen aus 1 mm dickem Feinsilberblech, das auf ein gleichgestaltetes Trolitul-Rähmchen aufgeschraubt ist. Zwei gegenüberliegende Seiten des Ag-Rähmchens sind nachträglich in $\frac{1}{3}$ mm Breite durchschnitten, so daß 2 mit den Schenkeln sich gegenüberstehende U-förmige Elektroden entstehen.

Ursprünglich wurden Messing-Elektroden verwandt. Die Silber-Elektroden bewährten sich aber besser. Sie ergaben ein wesentlich niedrigeres Widerstandsrauschen der Bolometer, doch blieb dieser Kontakt Bi—Ag immer ein heikler Punkt.

Zwei solche Elektroden-Rahmen sind vorn und hinten auf einer Metallplatte befestigt. Sie dient als Reflektor des von der Schicht durchgelassenen Strahlungs-Anteils und gleichzeitig zur Befestigung der beiden Schichten im Bolometer-Gehäuse. Der Abstand der Bolometerschicht vom Reflektor beträgt 2,5 mm.

Das Gehäuse des Normaltyps besteht aus drei Teilen: einem zylindrischen Mittelstück von 2 cm Höhe und 2,9 cm Durchmesser aus Aluminium, einem Boden, der das Doppel-Bolometer trägt und einem gegenüberliegenden Ultrarot-Fenster. Das Mittelstück besitzt einen Stutzen, in welchem ein Glasrohr mit Siegelack eingekittet ist. Durch dieses Rohr wird schließlich der erwünschte Gasdruck von etwa 1 Torr Luft eingestellt. Das Glasrohr wird dann abgeschmolzen.

Als Ultrarotfenster dient NaCl oder KRS 5 (Tallium-bromid-jodid-Mischkristall) in 3 mm Dicke. Die zwei Endflächen des Gehäuse-Mittelstückes, auf welchen das Fenster bzw. der Gehäuseboden aufliegt, sind sorgfältig plan geschliffen. Sie werden mit Apiezonfett gedichtet. Je ein vorn und hinten aufschraubbarer Überwurfing sorgt für festen Sitz des Fensters und des Gehäusebodens.

§ 4. Herstellung der Zelluloid-Membran und der Schicht

Als Trägerschicht für die Bolometer dienten Zelluloid-Membranen, die aus Zaponlack hergestellt wurden. Die von uns verwandte Herstellungsmethode ist ausführlich beschrieben in anderem Zusammenhang in der Z. Physik 108, 89 (1938).

Um eine rasche, angenäherte Dickenmessung der Membranen zu ermöglichen, wurde ein Satz Membranen etwa des Bereiches $0,03$ bis $0,2\mu$ hergestellt. Deren Dicke wurde optisch bestimmt durch Ausmessen ihres Reflexions-Vermögens für monochromatisches Licht mit Hilfe eines Pulfrich-Photometers. Durch rein visuellen Vergleich mit diesem Standard-Satz wurde dann die Dicke später hergestellter Membranen bestimmt. Der Zusammenhang zwischen Dicke und Farbe war etwa folgender: $0,13\mu$ strohgelb, $0,10\mu$ weiß, $0,08\mu$ eisengrau, $0,06\mu$ zart dunkelblau bei nahezu senkrechter Beleuchtung. Unterhalb $0,06\mu$ immer dunkler werdende Töne von schwärzlichem Blau. Schichten von $0,04\mu$ ließen sich noch gut für Bolometer verwenden, solche von $0,03\mu$ nur mit besonderer Vorsicht.

Das Aufkleben der Membranen auf die Ag-Elektroden erfolgte mit Spiritus. Wenn man die noch in ihrem ursprünglichen Rahmen gehaltene Membran auf den

Elektroden-Rahmen aufliegt und dann mit einem kleinen Haarpinsel andrückt, der mit Spiritus getränkt ist, so klebt die Membran sofort sehr haltbar auf dem Metall fest. Außerhalb des Elektroden-Rahmens erweicht die Membran durch den Spiritus so, daß man sie auf diese Weise mit dem Pinsel heraus schneiden kann.

Mitunter ist die Membran nach dem Aufziehen nicht vollkommen glatt, sondern verrät durch eine leichte Wellung eine ungleichmäßige Spannungs-Verteilung. Diesen Defekt kann man durch eine Nachbehandlung mit Spiritus-Dampf beheben. Man bringt dazu das fertig bespannte Bolometer-Rähmchen kurz über ein Schälchen mit Spiritus in 1 cm Entfernung. Die Membran wird dann zunächst gleichmäßig runzelig, nach Wegnahme aus dem Dampf strafft sie sich aber und wird einwandfrei faltenlos.

Auch der umgekehrte Fall kann eintreten, daß nämlich die Membran zu scharf gespannt ist und dann beim Einbringen ins Vakuum oder beim Bedampfen reißt. Als Gegenmaßnahme wurde folgende Anordnung beim Aufkleben der Membran auf den Elektrodenrahmen getroffen: Der Rahmen wurde auf ein ebenes Stück Blech gelegt, das in der Mitte eine kleine Öffnung hatte, durch die die Luft mit Hilfe einer Wasserstrahlpumpe in geringem Maße abgesaugt werden konnte. Dann ist die Membran etwas nach innen gezogen, wenn sie mit Spiritus aufgeklebt wird. Außerdem wird durch die Absaug-Öffnung überschüssiger Spiritus sofort entfernt.

Die Bi-Schicht wurde im Hochvakuum aufgedampft. Dazu wurde der Elektroden-Rahmen mit der Membran etwa 10 cm über dem zu verdampfenden Bi auf eine isolierende Fläche mit einer passend gestalteten Öffnung gesetzt. Sie diente als Schablone. Da bei dem Aufziehen der Membran die ganze Oberfläche der Ag-Elektroden mit der isolierenden Membran überklebt war, mußte zuvor die Membran bis auf einen etwa $\frac{1}{4}$ mm breiten Streifen von den Elektroden-Flächen mit einem Messer weggeschabt werden. Die Ag-Schicht der Elektroden wurde dabei auch noch mit dem Messer durch einige Kratzer aufgeraut, um eine größere, frische Kontaktfläche für das aufgedampfte Bi zu bilden. Bei dem Bedampfungs-Prozeß wurde durch die Schablone hindurch zunächst der Streifen, der als Bolometer-schicht dienen sollte, und gleichzeitig ein breites Gebiet der Elektroden bedampft. Dabei lag eine Spannung von 0,1 Volt an den Elektroden, so daß die Bedampfung unterbrochen werden konnte, wenn der gewünschte Widerstands-Wert erreicht war. Die Bedampfung wurde bei einem Druck von etwa 10^{-5} Torr in etwa 30 Sekunden durchgeführt. Zur Verbesserung des Kontaktes an den Elektroden wurde dann sofort anschließend eine weitere Schicht Bi auf die Elektroden und einen ganz schmalen, angrenzenden Teil der Membran aufgedampft, während der eigentliche Bolometer-Streifen durch ein Metallstück abgedeckt war, das mit Hilfe eines Magneten im Hochvakuum vor den Mittelteil der Schablonen-Öffnung geklappt werden konnte.

Normalerweise wurden zwei Bolometer-Typen hergestellt, die eine mit einer Schichtfläche $7 \cdot 6 \text{ mm}^2$, die andere mit $2 \cdot 3 \text{ mm}^2$. Einige Exemplare wurden auch in der Größe $25 \cdot 25 \text{ mm}^2$ angefertigt. Die Zelluloid-Membran war meist $0,04 \mu$ dick. Der Quadratwiderstand der Schichten lag zwischen 100 und 300 Ohm.

§ 5. Die Übertemperatur der Schicht im Betrieb

Um einen Einblick in die Betriebsbedingungen der Bolometer zu bekommen, wurden einige Untersuchungen durchgeführt, über die im folgenden berichtet werden soll.

Eine Bolometer-Membran nimmt unter normalen Betriebsbedingungen durch den Meß-Strom eine erhebliche Übertemperatur an. Diese darf als gleichmäßig

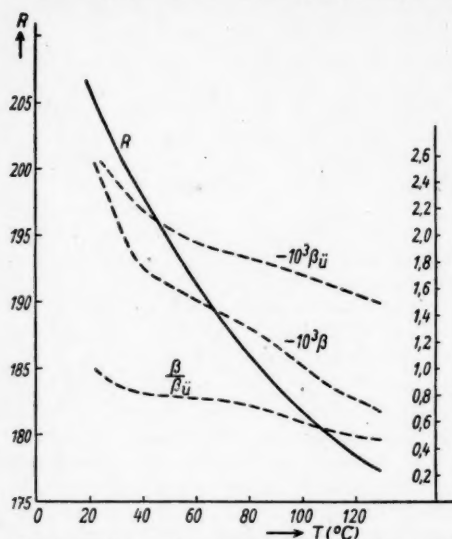


Abb. 2. Widerstand R einer Bi-Folie in Abhängigkeit von ihrer Temperatur (linke Ordinaten-Teilung). β und β_u sind der daraus abgeleitete differentielle Temperatur-Koeffizient und der Differenzen-Temperatur-Koeffizient, letzterer bezogen auf $T_0 = 20^\circ\text{C}$ und $R_0 = 206,4 \Omega$. Dazu die rechte Ordinaten-Teilung

bracht und dabei ihr Widerstand gemessen. Der Strom wurde so schwach gehalten, daß die Stromwärme zu vernachlässigen war. In der Abb. 2 ist das Ergebnis dargestellt. Daraus ist die Kurve der

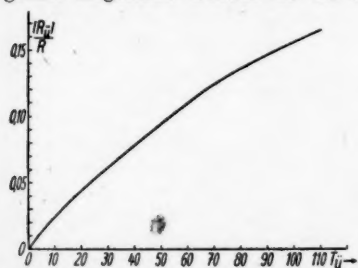


Abb. 3. Ordinate: Absoluter Betrag der relativen Widerstands-Änderung $|R - R_0|/R = R_u/R$. Abszisse: Übertemperatur $T_u = T - T_0$. $R_0 = 206,4 \Omega$, $T_0 = 20^\circ\text{C}$

über die ganze Fläche hin angesehen werden, da die Querleitfähigkeit der Membran klein ist gegenüber dem Wärme-Verluststrom senkrecht zur Membran. Es geht dies besonders direkt aus Beobachtungen hervor, die bei anderer Gelegenheit, nämlich bei evaporographischen Versuchen gemacht wurden⁴⁾. Dort war eine dünne Aluminium-Schicht als Absorber auf eine Zelluloid-Membran aufgedampft. Es zeigte sich, daß damit Bestrahlungs-Unterschiede auf der Membran bis zu Entfernungen von wenigen Zehntel Millimetern herab nachgewiesen werden konnten.

Um bei den folgenden Betrachtungen aus der Widerstandsänderung der Bolometer-Membran bei einer Erwärmung auf ihre Temperatur rückschließen zu können, wurde eine Membran in einem geschlossenen Glasgefäß unter vermindertem Druck in einem Ölbad auf verschiedene Temperaturen ge-

$$(R - R_0)/R = R_u/R \quad (3)$$

in Abb. 3 abgeleitet. Diese Kurve diente als Eichkurve, um auch bei andern Bolometer-Schichten, soweit sie im Quadrat-Widerstand nicht gar zu abweichend waren, aus der relativen Widerstands-Änderung die Übertemperatur zu bestimmen.

Aus obiger Messung wurde ferner abgeleitet der differentielle Temperaturkoeffizient

$$\beta = \frac{1}{R} \frac{dR}{dT} \quad (4)$$

⁴⁾ M. Czerny u. P. Mollet, Z. Physik 108, 99 (1938).

und der Temperatur-Koeffizient für die Differenz gegen Zimmer-Temperatur

$$\beta_a = \frac{1}{R} \frac{R_a}{T_a} \quad (5)$$

Beide Werte sind in Abb. 2 eingetragen.

Der hier gefundene Wert des Temperatur-Koeffizienten ist dem Betrage nach knapp halb so groß wie der von dicken Bi-Schichten und weist das entgegengesetzte Vorzeichen auf. Ein negativer Temperatur-Koeffizient ist für Bolometer als günstig anzusprechen, da dann bei einer Bestrahlung der Widerstand abnimmt und daher der Brückenstrom die Bolometer-Folie etwas stärker heizt, falls der Spannungsabfall an der Folie konstant gehalten wird. Es tritt also ein kleiner Effekt in Richtung einer Verstärker-Wirkung auf. Daß dagegen der absolute Betrag des Temperatur-Koeffizienten bei den dünnen Folien mit wachsender Temperatur abnimmt, ist natürlich ungünstig für das Bolometer.

Bei einer mittleren Belastung des Bolometers kann man etwa mit den Werten $\beta = -1,6 \cdot 10^{-3}$, $\beta_a = -2 \cdot 10^{-3} \text{ Grad}^{-1}$, $\beta/\beta_a = 0,8$ rechnen. — Temperaturen höher als 130° C vertragen die Zelluloid-Membranen nicht.

Über das Verhalten der Bi-Bolometer-Folien, wenn sie durch einen elektrischen Strom geheizt werden, geben Meßresultate einen Überblick, die in Tabelle 1 und 2 zusammengestellt sind. Bei den Messungen der Tabelle 1 befand sich die Schicht in Luft von 1,5 mm Hg-Säule, wie es im praktischen Betrieb auch meist der Fall

Tabelle 1
 $p = 1,5 \text{ Torr Luft}$ $R_0 = 195,15 \Omega$ $F = 0,6 \cdot 0,7 \text{ cm}^2$

V (Volt)	R (Ω)	$ R_a $ (Ω)	$\frac{ R_a }{R}$ (in 10^{-2})	T_a ($^\circ \text{ C}$)	$L = \frac{V^2}{R F}$ (in $10^{-3} \frac{\text{Watt}}{\text{cm}^2}$)	$W_e = \frac{L}{T_a}$ (in $10^{-4} \frac{\text{Watt}}{\text{cm}^2 \text{ Grad}}$)
0,25	195,1	0,05	0,026	0,15	0,76	50,7
1,25	193,6	1,55	0,80	3,2	19,2	60,1
2,25	189,8	5,35	2,82	11,9	63,5	53,4
3,25	184,8	10,35	5,60	27	136	50,4
4,25	178,8	16,30	9,11	48	240	50,1
5,25	173,0	22,15	12,80	73	379	51,9
6,25	168,5	26,65	15,80	102	552	54,0

Tabelle 2
 $p = 10^{-3} \text{ Torr}$ $R_0 = 185,9 \Omega$ $F = 0,6 \times 0,7 \text{ cm}^2$

V (Volt)	R (Ω)	$ R_a $ (Ω)	$\frac{R_a}{R}$ (in 10^{-2})	T_a ($^\circ \text{ C}$)	$L = \frac{V^2}{R F}$ (in $10^{-3} \frac{\text{Watt}}{\text{cm}^2}$)	$W_e = \frac{L}{T_a}$ (in $10^{-4} \frac{\text{Watt}}{\text{cm}^2 \text{ Grad}}$)
0,1	185,8	0,1	0,054	0,3	0,13	4,27
0,5	184,1	1,7	0,90	3,5	3,23	7,97
1,0	180,7	5,1	2,82	11,9	13,2	11,1
1,5	175,8	10,0	5,69	27,4	30,5	11,1
2,0	170,3	15,5	9,10	47,8	55,9	11,7
2,5	165,7	20,1	12,1	68	89,8	13,2

war. Bei Tabelle 2 war der Druck 10^{-3} mm Hg-Säule, um das unterschiedliche Verhalten im Hochvakuum hervortreten zu lassen.

In den Tabellen bedeutet V die an das Bolometer angelegte Spannung, R_u die durch diese Strombelastung eingetretene Widerstandsänderung, T_u die mit Hilfe von Abb. 3 daraus berechnete Übertemperatur der Bolometer-Folie.

$$L = \frac{V^2}{RF} \quad (6)$$

ist die pro cm^2 der Schicht zugeführte elektrische Heizleistung und daher auch gleichzeitig die Stärke des Wärme-Verluststromes pro cm^2 .

$$W_v = \frac{L}{T_u} \quad (7)$$

ist die Stärke des Wärmeverluststromes pro cm^2 und pro Grad Übertemperatur.

Ein Vergleich der beiden letzten Spalten der Tabellen 1 und 2 zeigt, daß W_v die Wärmeverluststromstärke pro Grad Übertemperatur, erstens bei 1,5 Torr erheblich größer ist als bei 10^{-3} Torr, daß sie ferner bei 1,5 Torr fast konstant ist, was die folgenden Betrachtungen sehr erleichtert, während sie bei 10^{-3} Torr mit wachsender Belastung ansteigt. Das beweist, daß der Wärmeverlust bei dem höheren Druck im wesentlichen auf dem Wege der Wärmeleitung und Konvektion im Gase erfolgt, dagegen bei dem niedrigeren Druck auch die Wärmestrahlung bedeutsam wird. Denn nach dem Stefan-Boltzmannschen Gesetz steigt die Wärmestrahlung stärker als proportional T_u an. Wäre allerdings die Strahlung bei dem niedrigen Druck die einzige Ursache des Wärmeverlustes, so dürfte die letzte Zahl der letzten Spalte der Tabelle 2 höchstens 8,3 statt 13,2 betragen.

§ 6. Das Verhalten der Schicht bei Einstrahlung

Das unbestrahlte Bolometer nimmt die Übertemperatur T_u von etwa 30 oder mehr Grad C an, wenn die normale konstante Betriebsspannung am Bolometer liegt. Wenn dem Bolometer außerdem noch zusätzlich Energie in Form von Strahlung zugeführt wird, so erhöht sich seine Temperatur um den weiteren Betrag ΔT , der in der Praxis nur einen kleinen Bruchteil eines Grades bis herab zum millionstel Grad C beträgt. Der einfachste Ansatz für den zeitlichen Verlauf von ΔT in Abhängigkeit von der zeitlichen Veränderung der Bestrahlung ist

$$\begin{aligned} W_k \frac{d\Delta T}{dt} + W_v \Delta T &= 0 & t < 0 \\ &= Q \frac{1}{2} (1 + \cos \omega t) & t \geq 0. \end{aligned} \quad (8)$$

Darin soll W_k der Betrag der Wärmekapazität des Bolometers sein, der auf 1 cm^2 der Bolometerfläche entfällt. Diese in Wattsec/ cm^2 zu messende Größe setzt sich zusammen aus der Wärmekapazität der eigentlichen Bolometerschicht und der durchaus nicht zu vernachlässigenden Wärmekapazität des umgebenden Gases, soweit dies mit erwärmt wird. Zur Vereinfachung soll angenommen werden, daß W_k einen frequenzunabhängigen Wert besitzt, was in Wirklichkeit nicht genau zutrifft, da bei einer wechselnden Bestrahlungsstärke sich Wärmewellen in der Umgebung ausbilden, die besonders in dem umgebenden Gas, das in thermischer Strömung begriffen ist, wohl nur sehr schwer in Rechnung zu setzen wären. Bei ruhendem Gas wäre die Eindringtiefe der Wärmewellen proportional zu $\sqrt{\frac{\lambda T}{p v}}$.

λ kennzeichnet die Wärmeleitung des Gases, p seinen Druck, ν die Frequenz der modulierten Bestrahlung. Das Glied $W_v \Delta T$ soll den zusätzlichen Wärmeverluststrom pro Flächeneinheit der Schicht bedeuten, der dadurch entsteht, daß das Bolometer eine zusätzliche Temperaturerhöhung ΔT über seine schon bestehende Übertemperatur T_u hinaus besitzt. Unter Berufung auf die Konstanz der Werte von $W_v = L/T_u$ in der letzten Spalte der Tabelle 2 wird bei Bolometern mit geringer Gasfüllung für W_v in der obigen Gleichung der gleiche Wert angenommen, den man nach dem Vorbild der Tabelle 2 aus dem Verhältnis L/T_u berechnet. Bei Vakuumbolometern wäre diese Vereinfachung nicht möglich.

Der Ausdruck $Q \frac{1}{2} (1 + \cos \omega t)$ soll den zeitlichen Verlauf der Energiezufuhr in Form von Strahlung beschreiben. Er ist auf 1 cm^2 der Bolometerfläche bezogen. Es ist also eine mit der Frequenz $\omega = 2\pi\nu$ sinusförmig modulierte Bestrahlung⁵⁾ des Bolometers angenommen.

Wenn die Strahlenquelle ein schwarzer Körper der Temperatur T_s und der Strahlenfläche F_s ist, T_z die Zimmertemperatur bedeutet, A_R das Absorptionsvermögen der Bolometerfolie für diese Strahlung und r den Abstand zwischen schwarzem Strahler und Bolometer, so ist

$$Q = A_R I = A_R \frac{\sigma F_s}{\pi r^2} (T_s^4 - T_z^4) \quad \text{Watt cm}^{-2}. \quad (9)$$

Auch in dem Glied $Q \frac{1}{2} (1 + \cos \omega t)$ steckt eine gewisse Vereinfachung gegenüber den wirklichen Verhältnissen, insofern bei dem Bolometer die Temperaturänderung eine Widerstandsänderung und damit im allgemeinen eine kleine Änderung ΔL der Heizleistung mit sich bringt. Denkt man sich die Spannung am Bolometer konstant gehalten, so ist

$$\Delta L = -L \frac{\Delta R}{R} = -L \beta \Delta T.$$

Diese Änderung der Heizleistung kann man in das Wärmeverlust-Glied auf der linken Seite der Gl. (8) einbeziehen und dort genauer statt W_v die Größe $W'_v = W_v + L\beta = L \left(\frac{1}{T_u} + \beta \right)$ setzen. Man erhält z. B. $W'_v = 0,95 W_v$ für $\beta = -1,6 \cdot 10^{-3} \text{ grad}^{-1}$ und $T_u = 30^\circ$.

Die Integration der Differentialgleichung (8) unter der Anfangsbedingung $\Delta T = 0$ für $t = 0$ ergibt für die Temperaturerhöhung

$$\Delta T = \frac{Q}{2 W'_v} \left(1 - \frac{2 W'_v + \omega^2 W_k^2}{W'_v + \omega^2 W_k^2} e^{-\frac{W'_v}{W_k} t} \right) + \frac{Q}{2} \frac{1}{\sqrt{W'_v + \omega^2 W_k^2}} \cos(\omega t - \varphi), \quad (10)$$

worin die Phasenverschiebung $\varphi = \arctg \frac{\omega W_k}{W'_v}$ ist. Wir betrachten zwei Grenzfälle:

1. Den rein periodischen Fall. Wenn nach hinreichender Zeit das Glied mit der Exponentialfunktion abgeklungen ist, findet man eine periodische Temperaturschwankung mit der Amplitude

$$\Theta(\omega) = \frac{Q}{2} \frac{1}{\sqrt{W'_v + \omega^2 W_k^2}} \quad (11)$$

⁵⁾ W. Kofink, Lichtmodulation durch Lochscheiben und Drehflügel, Das Licht 14, 108 (1944).

über eine mittlere Temperaturerhöhung $\frac{Q}{2W_v}$ überlagert. Daraus ergibt sich die „Halbwertsfrequenz“ ν_1 , d. h. die Frequenz, bei welcher die Amplitude auf die Hälfte derjenigen bei $\nu = 0$ herabgesunken ist zu

$$\nu_1 = \frac{\sqrt{3} W_v}{2\pi W_k}. \quad (12)$$

Als Beleg dafür, inwieweit der Ausdruck (11) trotz der im Ansatz gemachten Vereinfachungen die experimentellen Befunde wiederzugeben gestattet, diene die

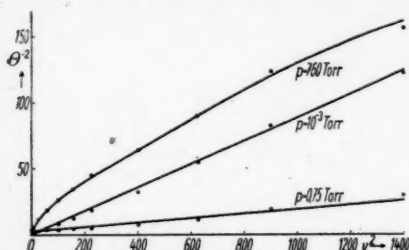


Abb. 4. Zusammenhang zwischen der Temperaturschwankungs-Amplitude Θ eines Bolometers (in willkürlichen Einheiten aufgetragen) und der Frequenz ν , mit der die einfallende Strahlung moduliert ist

Abb. 4. Nach (11) sollte $\Theta(\omega)^2$ linear abhängig sein von ν^2 . Man erkennt, daß bei den beiden niedrigen Gasdrucken, wie sie im Folgenden im wesentlichen in Betracht kommen, die Meßpunkte bis zu einer Frequenz von 35 sec^{-1} hinreichend auf einer Geraden liegen. Die Bolometerschicht hatte 123Ω Widerstand und war auf einer Membran von $0,12 \mu$ aufgetragen.

2. Im Falle konstanter Einstrahlung, d. h. $\omega = 0$, ergibt sich für den zeitlichen Verlauf des Temperaturanstieges

$$\Delta T_0 = \frac{Q}{W_v} \left(1 - e^{-\frac{W_v}{W_k} t} \right). \quad (13)$$

Der Endwert ergibt sich daraus

$$\Delta T_{0/t \rightarrow \infty} = \frac{Q}{W_v}. \quad (14)$$

Definiert man als Einstellzeit τ_e die Zeit, nach welcher die Temperaturerhöhung bis auf $e^{-2\pi} \approx 0,2\%$ ihren Endwert erreicht hat⁶⁾, so wird

$$\tau_e = 2\pi \frac{W_k}{W_v}. \quad (15)$$

Zwischen der Einstellzeit und der Halbwertsfrequenz besteht daher die Beziehung

$$\tau_e \nu_1 = \sqrt{3}. \quad (16)$$

Aus diesen Formeln ersieht man, daß das Verhalten des Bolometers im wesentlichen durch die zwei Größen W_v und W_k bestimmt ist. W_v ist nach den im vorigen Abschnitt gegebenen Beziehungen der Messung zugänglich, denn es ist

$$W_v = \frac{L}{T_a} = \frac{V^2}{R F} \frac{1}{T_a} = \frac{V^2}{R F} \frac{\beta_a R}{R_a}. \quad (17)$$

Aus W_v kann man W_k ableiten nach den Formeln (12) oder (16), je nachdem, ob man die Halbwertsfrequenz oder die Einstellzeit des Bolometers gemessen hat.

⁶⁾ Diese Definition der Einstellzeit ist die gleiche, wie man sie für die Einstellzeit eines kritisch aperiodisch gedämpften Galvanometers macht, wenn man festsetzt, daß dessen Einstellzeit gleich seiner vollen Schwingungsdauer in ungedämpften Falle sei.

Zum Falle konstanter Einstrahlung sei noch angeführt: Formel (14) gibt die bei konstanter Einstrahlung sich einstellende Übertemperatur $\Delta T_{0/t \rightarrow \infty}$. Damit steht im Zusammenhang die relative Widerstandsänderung:

$$\frac{\Delta R}{R} = \beta \Delta T_{0/t \rightarrow \infty} \quad (18)$$

$$\frac{\Delta R}{R} = \beta \frac{Q}{W_v} = \frac{\beta A_R I}{W_v} = \frac{\beta T_a A_R I}{L} = A_R \frac{\beta}{\beta_a} \frac{R_a}{R L} I.$$

Eine Leistungsgrenze des Bolometers als Strahlungsempfänger ist durch das Widerstandsrauschen des Bolometers gegeben. Im günstigsten Falle hat dieses den aus der Nyquist-Formel⁷⁾ folgenden Betrag

$$\bar{V}_N = \sqrt{4 k T R \Delta \nu} \quad (\text{Volt}). \quad (19)$$

Dabei ist $k = 1,37 \cdot 10^{-23}$ Wattsec Grad⁻¹ die Boltzmannsche Konstante und $\Delta \nu$ die Breite des mitwirkenden Frequenzbandes. Als kleinste meßbare Strahlungsintensität I_0 wird man die ansprechen, die am Bolometer eine Spannungsänderung ΔV hervorruft, die z mal größer als \bar{V}_N ist, wobei man etwa $z = 3$ wählen wird. Daraus ergibt sich zunächst die kleinste nachweisbare Temperatur-Änderung des Bolometers durch die Beziehung

$$z \bar{V}_N = \Delta V = i \Delta R = \frac{V}{R} \Delta R = V \beta \theta. \quad (20)$$

Wird sinusförmig modulierte Bestrahlung des Bolometers vorausgesetzt, so folgt aus (11), (12) und (20) die kleinste noch nachweisbare Strahlungs-Intensität zu

$$I_0 = \frac{4 z W_v}{A_R V |\beta|} \sqrt{k T R \Delta \nu \left(1 + 3 \left(\frac{v}{v_t}\right)^2\right)} \quad \text{Watt cm}^{-2}. \quad (21)$$

§ 7. Empfindlichkeitsmessungen bei konstanter Einstrahlung

Bei konstanter Einstrahlung soll nach Formel (18) der theoret. Betrachtungen des vorigen Paragraphen die relative Widerstandsänderung eines Bolometers bei einer Bestrahlung mit der konstanten Intensität I um so größer sein, je größer $A_R \beta T_a / L$ ist. Es wurde daher eine Meßreihe durchgeführt mit dem Ziel, die Größe dieses Ausdruckes bei Bi-Schichten verschiedener Dicke oder — was auf das gleiche herauskommt — verschiedenem Quadrat-Widerstand direkt zu messen. Dazu wurde folgendermaßen vorgegangen: Die Bedampfung der Zelluloid-Membran mit Wismut im Vakuum wurde schrittweise mit Unterbrechungen durchgeführt. Jedesmal während einer Unterbrechung wurde mit Hilfe einer Wheatstoneschen Brücke zweierlei gemessen: 1. Der Widerstand R der Bi-Schicht und 2. seine Änderung ΔR , hervorgerufen durch eine Bestrahlung der Schicht. Die Strahlung kam von einem schwarzen Körper von 300° C und fiel durch ein NaCl-Fenster in die Vakuum-Apparatur. Am Ort der Bi-Schicht betrug die Strahlungsdichte $I = 1,16 \cdot 10^{-3}$ Watt/cm². Die Schicht war 1 cm² groß. Die Spannung der Stromquelle der Wheatstoneschen Brücke wurde jeweils so einreguliert, daß die Heizleistung des Brückenstromes in der Schicht den konstanten Betrag $L = 5 \cdot 10^{-3}$ Watt/cm² besaß.

⁷⁾ H. Nyquist, Physic. Rev. **32**, 110 (1928); J. B. Johnson, Physic. Rev. **32**, 97 (1928); Müller-Pouillet, Lehrbuch der Physik, Bd. IV, 4. Teil, S. 279.

Etwa 3 mm unter der Schicht war ein Strahlen-Reflektor angebracht, der durch eine magnetische Vorrichtung ein- und ausgeklappt werden konnte. Dadurch konnte die Erhöhung der Strahlen-Absorption durch einen Reflektor bei den verschiedenen Schichtdicken messend verfolgt werden^{*)}.

In Tabelle 3 sind die Versuchsergebnisse zusammengestellt. Die Spalten 1 und 2 enthalten den Schichtwiderstand und den Spannungsabfall an der Bi-Schicht. In der dritten und vierten Spalte findet man den Ausschlag des Galvanometers im Nullzweig der Brücke beim Einschalten der Bestrahlung mit bzw. ohne Reflektor. Die Spalten 5 und 6 geben einen Vergleich zwischen dem Verhältnis dieser Galv.-Ausschläge und dem in § 1 theoretisch berechneten Verhältnis des Absorptionsvermögens A_R zu A , d. h. wieder mit und ohne Reflektor. Unter der Voraussetzung, daß die Übertemperatur durch den Reflektor nicht wesentlich verändert wird, d. h. daß βT_u beim Einklappen des Reflektors konstant bleibt, sollten beide Spalten übereinstimmen.

Tabelle 3

Heizleistung $L = \frac{1}{200} \text{ Watt cm}^{-2}$ = Const		Brückenstrom i_g mit ohne Reflektor		exp. i_g (mit) i_g (ohne)	theor. $\frac{A_R}{A}$	$\frac{ \Delta R }{R}$ (in 10^{-3})	$\frac{ \Delta R L}{RI}$ $= A_R \beta T_u$ (in 10^{-3})
$R (\Omega)$	$V (\text{Volt})$						
604	1,75	23	15	1,53	1,62	0,96	4,1
450	1,50	66	45	1,47	1,55	2,52	10,8
384	1,40	78	56	1,39	1,51	2,86	12,3
315	1,25	104	74	1,40	1,46	3,64	15,7
287	1,20	121	86	1,40	1,43	4,18	18,0
255	1,13	138	98	1,41	1,40	4,67	20,1
227	1,06	149	108	1,38	1,38	4,98	21,4
202	1,00	138	100	1,38	1,35	4,52	19,5
181	0,95	132	98	1,35	1,32	4,31	18,6
143	0,84	129	98	1,32	1,28	4,17	18,0
101	0,71	109	87	1,24	1,21	3,52	15,2
80	0,63	93	76	1,22	1,18	3,12	13,4
40	0,45	64	56	1,14	1,10	2,55	13,9
20	0,32	66	60	1,10	1,05	3,22	13,9

Aus der letzten Spalte der Tabelle ist der Verlauf des in Frage stehenden Ausdrucks $A_R \beta T_u / L$ als Funktion des Quadrat-Widerstandes der Schicht zu ersehen. Die Zahlenwerte durchlaufen um den Quadrat-Widerstand von 220 Ohm herum ein flaches Maximum. Das Maximum liegt also in der Nähe des Quadrat-Widerstandeswertes, für den A_R nach den theoret. Berechnungen des § 1 für sich allein schon einen Maximalwert annehmen soll.

Nebenbei ergab dieser Versuch auch einen Einblick in das Anwachsen des Kontaktes zwischen Folie und Silber-Rahmen und der Stabilität des Widerstandes der Schicht. Bei dünnen Schichten mit einem Quadrat-Widerstand über 500 Ohm herrschte bei der Widerstandsmessung in der Brücke eine beträchtliche Unruhe

^{*)} Das Einklappen des Reflektors rief schon ohne Bestrahlung einen geringen Ausschlag des Nullinstrumentes hervor, weil durch den Reflektor der Wärmeverluststrom der Membran verkleinert wurde. Der Effekt war aber gering und wurde bei der Auswertung berücksichtigt.

des Lichtzeigers des Galvanometers. Unter 400 Ohm etwa ist sie verschwunden. Man beobachtet nur noch eine stetige Widerstandszunahme ohne Unruhe vom Ende der Bedampfung ab. Auch sie verringert sich mit wachsender Schichtdicke.

§ 8. Verringerung der Einstellzeit

Die bisherigen Versuche bezogen sich sämtlich auf zeitlich konstante Einstrahlung; für sie spielte die Wärmekapazität des Bolometers keine Rolle. Betrachten wir nun die Wirkung zeitlich veränderlicher Einstrahlung. Nach (15) hängt die Einstellzeit $\tau_e = 2\pi \frac{W_z}{W_v}$ vom Verhältnis der Wärmekapazität zum Wärmeverluststrom ab. Man kann sie also auf 2 Weisen verringern: 1. durch Verkleinern der Wärmekapazität und 2. durch Erhöhung des Wärmeverluststromes. An Stelle der Einstellzeit bestimmen wir experimentell bequemer die Halbwertsfrequenz ν_1 , die nach Formel (16) der Einstellzeit umgekehrt proportional ist. Zu diesem Zweck fiel ein Strahlenbündel sichtbaren Lichts aus einer Nitralampe konstanter Heizstromstärke, durch eine mechanische Vorrichtung in der Intensität sinusförmig moduliert, auf die Bolometerschicht. Die Amplitude ihrer Temperaturänderung wurde ohne alle Verstärker-Einrichtungen direkt als Amplitude eines Torsionsfadengalvanometers gemessen, welches als Nullinstrument in der Brücke lag. Der bis etwa 30 Hz geringe, dann zunehmende Frequenzgang des Gal-

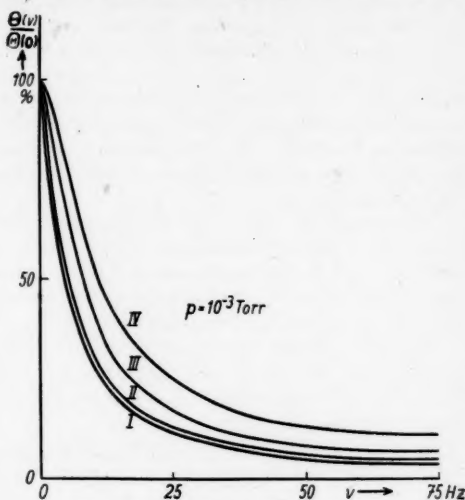


Abb. 5a

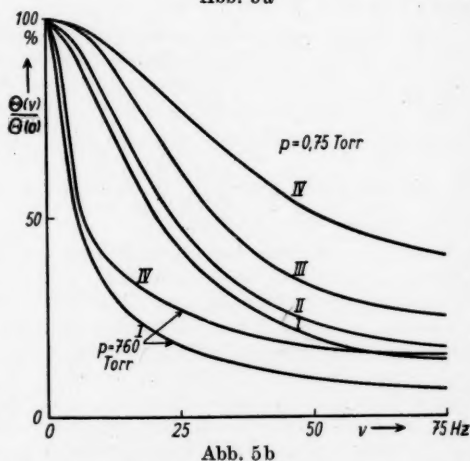


Abb. 5b

Abb. 5a und b. Zusammenhang zwischen der Frequenz ν einer sinusförmig modulierten Strahlung und der relativen Temperaturschwankungs-Amplitude $\Theta(\nu)/\Theta(0)$ bei 4 verschiedenen dicken Bolometern. Dicke der Zelluloidmembran bei I und II 0,12 μ , bei III und IV 0,04 μ . Widerstand der Wismut-Schicht bei I und III 160 Ohm, bei II und IV 320 Ohm

vanometers selbst wurde vorher mit einer trägheitslosen Photozelle aufgenommen und als Korrektur in Rechnung gestellt. Im Folgenden soll nun zuerst die Erhöhung der Halbwertsfrequenz durch eine Verkleinerung der Wärmekapazität W_k betrachtet werden. Sie kann erreicht werden durch eine Verringerung der Dicke des Metallbelages und der Membran. Im folgenden sind hierzu Meßresultate an 4 Bolometern von $0,7 \times 0,6 \text{ cm}^2$ Fläche mitgeteilt: 2 Bolometer, bezeichnet als I und II, mit relativ dicker Zelluloidmembran von $0,12 \mu$ und Metallschichten von 160 und 320Ω , und 2 Bolometer, bezeichnet mit III und IV, mit normaler Membran von $0,04 \mu$ Dicke und Metallschichten von 160 und 320Ω . Der Frequenzgang dieser Bolometer bei 3 verschiedenen Drucken ist in den Abb. 5a und 5b dargestellt.

Die Bolometer wurden zur Messung nacheinander in das gleiche Gehäuse eingebaut. Sie befanden sich also unter gleichen räumlichen Verhältnissen. Ferner wurde bei den Druckveränderungen die Heizspannung der Bolometer so verändert, daß der Bolometerwiderstand konstant blieb, was besagt, daß auch die Temperatur der Bolometerfolie konstant gehalten wurde.

Genauer gesagt: Bei Bestrahlung mit moduliertem Licht schwankt die Bolometertemperatur um den Wert, den sie bei Dauerbestrahlung mit der halben Intensität annehmen würde. Dieser Temperaturwert wurde bei der Druckvariation konstant gehalten.

Aus den Kurven geht hervor, daß die Verringerung der Dicke der Zelluloid-Membran die Wärmekapazität wirksamer herabsetzt als die Verringerung der Dicke der Bi-Schicht. Auch weitere Kontrollmessungen bei andern Drucken brachten das gleiche Resultat.

Eine kleine Abschätzung zeigt, daß dieses Resultat verständlich ist. Die Masse einer Bi-Schicht von 100 Ohm Quadrat-Widerstand ergab sich nach einer Wägung zu $m_1 = 7,8 \cdot 10^{-4} \text{ g/cm}^2$. Die spez. Wärme des Bi ist $c_1 = 0,12 \text{ Watt sec/g Grad}$. Also ist die Wärmekapazität $m_1 c_1 = 9,7 \cdot 10^{-6} \text{ Watt sec/cm}^2 \text{ Grad}$. Die Masse einer Zelluloid-Membran von $0,1 \mu$ Dicke ist $m_2 = 1,4 \cdot 10^{-4} \frac{\text{g}}{\text{cm}^2}$, die spezifische Wärme des Zelluloids ist $c_2 = 2,0 \frac{\text{Watt sec}}{\text{cm}^2 \text{ grad}}$, also die Wärmekapazität $m_2 c_2 = 2,8 \cdot 10^{-5} \frac{\text{Watt sec}}{\text{cm}^2 \text{ grad}}$ etwa das Dreifache von $m_1 c_1$.

Betrachten wir nun zweitens die Erhöhung der Halbwertsfrequenz durch die Vergrößerung des Wärmeverluststromes W_v . Dafür bestehen drei Möglichkeiten: durch Vergrößern des Gasdruckes, durch Verkleinern des Reflektorabstandes und durch Benutzung von Gasen höherer Wärmeleitfähigkeit. Man wird zunächst erwarten, daß eine Vergrößerung des Wärmeverluststromes die Empfindlichkeit des Bolometers so herabsetzt, daß ein solches Vorgehen praktisch nicht in Betracht gezogen werden kann. Das trifft aber nicht zu. Eine Erhöhung des Wärmeverluststromes gestattet nämlich das Anlegen höherer Heizspannungen an das Bolometer und dadurch in gewissen Grenzen eine Kompensation des Empfindlichkeitsverlustes. Ausgehend vom Hochvakuum, vergrößert eine Druckerhöhung bis zu einigen Torr merklich den Wärmeverlust, während der Beitrag des Gases zur Wärmekapazität gegenüber dem der Schicht noch nicht ins Gewicht fällt. Als obere Grenze für $m_g c_g$, die Wärmekapazität des Gases, bei unseren Bolometern mit 2,5 mm Reflektor- und etwa 1 cm Fensterabstand findet man bei $T_s = 290^\circ \text{ K}$ für $p = 1 \text{ Torr}$ den Wert $m_g c_g = 2 \cdot 10^{-6} \frac{\text{Watt sec}}{\text{cm}^2 \text{ grad}}$, wenn man die ganze Masse

des verfügbaren Gases als miterwärmt betrachtet. In diesem Druckbereich erhöht sich also die Halbwertsfrequenz mit wachsendem Druck (siehe Tabelle 4). Im Bereich von 50 Torr aufwärts dagegen liefert die miterwärmte Gasschicht einen immer merklicheren Beitrag zur Wärmekapazität, ohne daß der Wärmeverlust entsprechend anwächst. Die Wärmeleitfähigkeit und Zähigkeit des Gases erreichen im Bereich von einigen Torr ihren Endwert, so daß der Wärmeverlust durch Leitung und Konvektion nur noch wenig ansteigt. Ein Druck von einigen Torr wird daher ein Maximum der Halbwertsfrequenz erwarten lassen. Diese Zusammenhänge soll Tabelle 4 an einem Beispiel veranschaulichen. Es sind dort die Ergebnisse für das Bolometer IV zusammengestellt und ausgewertet. Die Übertemperatur über Zimmertemperatur wurde nach der Methode des § 5 zu $T_u = 51,4^\circ \text{C}$ bestimmt und bei der Druckvariation konstant gehalten durch Wahl einer passenden Betriebsspannung. Während die ersten 5 Spalten direkt dem Experiment entnommen sind, machen die 2 letzten von dem nicht in gleichem Maße gesicherten Wert für T_u Gebrauch. T_u ist jedoch dabei nur Proportionalitätskonstante; die relativen Werte von W_v und W_k als Funktionen des Druckes bleiben von der Ungenauigkeit von T_u unbeeinflusst.

Tabelle 4

Messungen am Bolometer IV, $0,04 \mu$ dicke Zelluloid-Schicht, 320Ω , $0,7 \times 0,6 \text{ cm}^2$ Fläche, bei allen Drucken auf 51°C Übertemperatur geheizt

Luftdruck p (Torr)	Betriebs- spannung V (Volt)	Heiz- leistung $L = W_v T_u$ (in 10^{-2} Watt cm^{-2})	Halb- werts- frequenz $\nu_{\frac{1}{2}}$ (sec $^{-1}$)	Wärmeverlust pro $^\circ \text{C}$ $\frac{L}{T_u}$ $W_v = \frac{L}{T_u}$ (in 10^{-4} Watt $\text{cm}^2 \text{ grad}$)	Wärme- kapazität W_k (in 10^{-6} Watt sec $\text{cm}^2 \text{ grad}$)
10^{-3}	1,51	2,19	11	4	11
0,015	2,03	3,97	16	8	13
0,03	2,34	5,27	20	10	14
0,11	3,30	10,48	37	20	15
0,25	3,55	12,13	43	24	15
0,75	3,91	14,72	53	29	15
1,4	4,04	15,71	55	31	15
10	4,18	16,82	56	33	16
75	4,20	17,00	29	33	32
760	4,30	17,80	7	35	136

Man sieht, wie der Wärmeverluststrom W_v vom Werte $4 \cdot 10^{-4}$ als Strahlungsverlust bei Hochvakuum rasch auf etwa $30 \cdot 10^{-4} \frac{\text{Watt}}{\text{cm}^2 \text{ grad}}$ bei einigen Torr ansteigt und von dort bis Atmosphärendruck nur noch wenig größer wird.

Zur Wärmekapazität W_k trägt bei Hochvakuum allein die Membran und die Schicht $m_1 c_1 + m_2 c_2 = 11 \cdot 10^{-6}$ bei; im Druckgebiet einiger Torr kommt dann ein kleiner Beitrag von etwa $4 \cdot 10^{-6}$ des Gases hinzu, der sich rasch auf etwa $125 \cdot 10^{-6} \frac{\text{Watt sec}}{\text{cm}^2 \text{ grad}}$ bei Atmosphärendruck vergrößert.

Der soeben in der Tabelle 4 dargelegte Gang der Halbwertsfrequenz bei einer Druckvariation (aber festgehaltener Heiztemperatur T_u) und die ebenfalls in der Tabelle zum Ausdruck gebrachte Möglichkeit, die Strombelastung und damit die

Empfindlichkeit des Bolometers bei wachsendem Druck zu erhöhen, gibt das Verständnis für die in der folgenden Abb. 6 dargestellten Messungsergebnisse.

Auf der Abszisse ist die Modulationsfrequenz der auffallenden Strahlung aufgetragen, auf der Ordinate die Ausschläge des Galvanometers in der Meßbrücke als Kennzeichen der Empfindlichkeit des Bolometers.

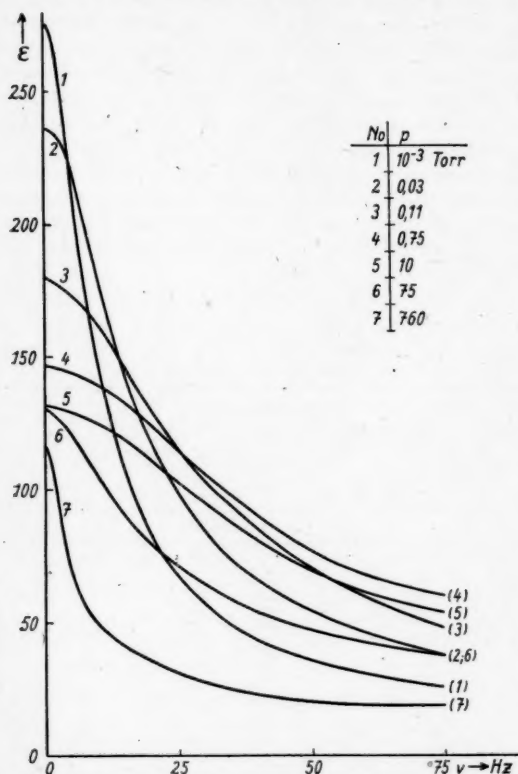


Abb. 6. Zusammenhang zwischen der Frequenz ν einer sinusförmig modulierten Strahlung und der Ausschlags-Amplitude ϵ des Nullinstrumentes bei einer Bolometer-Anordnung. Variiert wurde der Druck p im Bolometer, dagegen wurde T_a durch entsprechende Wahl der Heizspannung konstant gehalten

6 und 7 nimmt die Halbwertsfrequenz dagegen mit wachsendem Druck wieder ab und die Kurven können daher die andern nicht mehr so überschneiden, daß sie bei irgendeiner Frequenz zu maximaler Empfindlichkeit führen.

Für die Abnahme der Bolometer-Empfindlichkeit bei der Modulations-Frequenz $\omega = 0$ (in der Abb. 6 also für die Werte auf der Ordinaten-Achse) ergibt sich folgender theoretischer Zusammenhang mit der Größe W_v : Nach den Formeln 6

Die Intensität der Strahlung war bei allen Frequenzen die gleiche. Jede einzelne Kurve ist bei einem bestimmten Gasdruck aufgenommen, der in der Abbildung angegeben ist. Die Heizspannung am Bolometer wurde \bullet eingestellt, daß die Temperatur T_a konstant blieb.

Die Kurven der Abb. 6 zeigen Folgendes:

Für $\nu = 0$ nimmt die Empfindlichkeit monoton ab mit wachsendem Druck, wie man es erwarten wird, allerdings gemildert durch die Zunahme der Heizspannung.

Für höhere Frequenzen treten Überschneidungen der Kurven auf. Von der Kurve 1 bis zur Kurve 4 wird der Abfall der Empfindlichkeit mit wachsender Frequenz immer geringer entsprechend der mit dem Druck zunehmenden Halbwertsfrequenz. Infolgedessen erhält man das wesentliche Resultat, daß es für jede Frequenz einen bestimmten Druck gibt, der maximale Empfindlichkeit ergibt. Bei den Kurven 5,

und 7 ist $i_B = V/R = \sqrt{W_v F T_u/R}$. Daraus folgt zunächst, daß wenn T_u konstant gehalten wird, i_B und V proportional zu $\sqrt{W_v}$ wird. Nun liege das Bolometer in einer Wheatstoneschen Brücke. Es werde von dem Strom i_B durchflossen. Die Brücke sei zunächst abgeglichen, wenn das Bolometer unbestrahlt und sein Widerstand R ist. Erhöht man seine Temperatur um ΔT_0 durch eine Bestrahlung mit der Intensität I , so ändert sich R um ΔR und es fließt ein Strom i_G durch das Nullinstrument, der sich nach der Theorie der Brücke in der Form $i_G = -i_B \Delta R/W(R_v)$ schreiben läßt, wobei $W(R_v)$ ein Ausdruck ist, in dem die in der Brücke befindlichen Widerstände linear auftreten. Daher wird

$$\begin{aligned} i_G &= -i_B R \beta \Delta T_0 / W(R_v) = -i_B R \beta A_R I / W_v W(R_v) \\ &= -\sqrt{W_v F T_u / R} R \beta A_R I / W_v W(R_v) = -\sqrt{F T_u R / W_v} \beta A_R I / W(R_v) \\ i_G &= \text{const} \sqrt{1/W_v}. \end{aligned}$$

Nebenbei sei darauf hingewiesen, daß sich aus obiger Formel ergibt, daß für Bolometer, die bei verschiedener Fläche F gleiche Werte von R besitzen (also z. B. alle quadratische Fläche besitzen), die Empfindlichkeit proportional zu \sqrt{F} anwächst. Man kann leicht zeigen, daß diese hier für den statischen Fall abgeleitete Beziehung auch bei beliebigen Modulations-Frequenzen gilt.

§ 9. Beobachtung des veränderlichen Widerstandstemperaturkoeffizienten am Überschneiden von Frequenzkurven

Wenn man nicht die Temperatur wie im vorigen Paragraphen, sondern die Betriebsspannung konstant hält, verzichtet man auf die Empfindlichkeitssteigerung, welche die Erhöhung der Betriebsspannung bei wachsendem Druck mit sich bringt. Trotzdem tritt noch ein zwar kleiner, aber über der Fehlergrenze liegender Überschneidungseffekt der Frequenzkurven derselben Art wie in § 8 auf. Auch bei konstanter Betriebsspannung ist die Empfindlichkeit oberhalb einer gewissen Frequenz bei einem Füllgasdruck von etwa 1 Torr größer als diejenige bei Hochvakuum. Tabelle 5 zeigt diesen Sachverhalt an einem Bolometer von $7 \times 6 \text{ mm}^2$ Schichtfläche bei konstanter Betriebsspannung $V = 1,25 \text{ Volt}$ und dem Widerstand von 162Ω bei $p = 10^{-3}$, $172,5 \Omega$ bei $p = 0,5 \text{ Torr}$.

Tabelle 5

$\nu(\text{sec}^{-1})$	0	5	10	15	30	45	60
$p(\text{Torr})$							
10^{-3}	184	116	70	47	25	15	13
0,015	143	110	75	54	29	19	15
0,04	113	95	71	53	30	20	16
0,12	69	65	56	46	29	19	16
0,5	48	46	42	38	26	19	17

Die höchsten Empfindlichkeiten in jeder Spalte, d. h. bei konstanter Frequenz und wachsendem Druck, sind durch Fettdruck hervorgehoben; sie verlagern sich für höhere Frequenzen nach höheren Drucken, ein Verhalten, das dem in der Abb. 6 dargestellten analog ist.

Bei gleicher Betriebsspannung wird die Schicht im Hochvakuum heißer als bei Gasfüllung. Da für Bi-Schichten der Betrag des Widerstandstemperaturkoeffizienten $|\beta|$ bei Temperaturerhöhung abnimmt, kann die Empfindlichkeit bei Gasfüllung oberhalb einer bestimmten Frequenz ω_0 diejenige bei Hochvakuum übertreffen, obwohl sie bei $\omega = 0$ tiefer liegt.

Wir leiten die Bedingung dafür ab, wobei wir wieder Wärmeverlust und Wärmekapazität als frequenzunabhängig voraussetzen. Der Schnittpunkt ω_0 zweier Frequenzkurven ohne und mit Gasfüllung folgt bei konstanter Betriebsspannung aus

$$\frac{\beta_1}{\sqrt{W_v^2(0) + \omega_0^2 W_k^2(0)}} = \frac{\beta_2}{\sqrt{W_v^2(p) + \omega_0^2 W_k^2(p)}},$$

worin β_1 bzw. β_2 die Temperaturkoeffizienten des Schichtwiderstandes ohne und mit Gasfüllung, $W_v(0)$ und $W_k(0)$ bzw. $W_v(p)$ und $W_k(p)$ die entsprechenden Wärmeverluste und -kapazitäten bezeichnen. Man erhält

$$\omega_0^2 = \frac{\left(\frac{W_v(p)}{\beta_2}\right)^2 - \left(\frac{W_v(0)}{\beta_1}\right)^2}{\left(\frac{W_k(0)}{\beta_1}\right)^2 - \left(\frac{W_k(p)}{\beta_2}\right)^2}.$$

Nun folgt aus den Versuchen, daß für $\omega = 0$ immer das Hochvakuumbolometer die höchste Empfindlichkeit besitzt, d. h.

$$\frac{|\beta_1|}{W_v(0)} > \frac{|\beta_2|}{W_v(p)}.$$

Also ist der Zähler von ω_0^2 positiv; soll also ein Schnittpunkt ω_0 zustande kommen, so muß auch der Nenner positiv sein. Daraus folgt die Ungleichung

$$\frac{\beta_2}{\beta_1} > \frac{W_k(p)}{W_k(0)},$$

d. h. der Temperaturkoeffizient $|\beta|$ muß mit wachsendem Druck und der damit verbundenen Temperaturabnahme stärker zunehmen als die Wärmekapazität.

Die gleiche Überlegung kann man für die in § 8 gegebene andere Voraussetzung durchführen. An Stelle von β in den obigen Formeln hat man dann nur die Betriebsspannung V zu setzen; denn, während in § 9 V konstant und β veränderlich ist, war in § 8 β konstant und V veränderlich.

§ 10. Variation des Reflektorabstandes und des Füllgases

In § 8 wurden als zwei weitere Wege, die Einstellgeschwindigkeit zu erhöhen, die Verringerung des Reflektorabstandes und ein Füllgas höherer Wärmeleitfähigkeit erwähnt. Wir verringerten den Reflektorabstand auf 0,35 mm, d. h. auf etwa $\frac{1}{7}$ des normalen Abstandes von 2,5 mm. Statt der Halbwertsfrequenz $\nu_1 = 50 \text{ sec}^{-1}$ bei 2,5 mm ergab sich $\nu_1 = 86,5 \text{ sec}^{-1}$ bei 0,35 mm Reflektorabstand unter Luftfüllung von 0,7 Torr. Wenn man diese Bolometer mit Wasserstoff füllt, erzielt man eine weitere, ganz wesentliche Erhöhung der Halbwertsfrequenz, worüber Tabelle 6 berichtet:

Tabelle 6
 $F = 7 \times 6 \text{ mm}^2$; $a = 0,35 \text{ mm}$; $R = 223 \Omega$

p (Torr) H_2	0,45	0,85	1,6	3,5	42	500
ν_1 (sec^{-1})	157	230	445	319	283	149

Diese Halbwertsfrequenzen sind aus dem Verhältnis der gemessenen Amplituden bei $\nu = 75 \text{ sec}^{-1}$ und $\nu = 0$ extrapoliert worden; bei der Extrapolation wurde die Formel (11) mit frequenzunabhängigem W_v und W_k benutzt. Da die Wärmeleitfähigkeit des H_2 etwa 7mal größer ist als die der Luft, so ist die Eindringtiefe der Wärmewelle bei der gleichen Frequenz, Temperatur und Druck, für H_2 $\sqrt{7}$ mal größer als für Luft; oder der Frequenzbereich, in welchem die Eindringtiefe oberhalb eines festen Betrages (z. B. des Reflektorabstandes) bleibt, ist für H_2 7mal größer als für Luft unter gleichem Druck und gleicher Temperatur. Daher kann die Näherung frequenzunabhängiger Werte von W_v und W_k bei unserer Extrapolation noch benutzt werden.

Durch die Füllung mit Wasserstoff und die Verkleinerung des Reflektorabstandes wird der Wärmeverluststrom bei 1,6 Torr etwa auf das 9fache gegenüber dem Normalbolometer erhöht, wie sich aus den Experimenten ergibt. Dies bedeutet, daß dem Bolometer bei derselben Übertemperatur die neunfache Heizleistung zugeführt werden kann, d. h. die dreifache Betriebsspannung. Die statische Empfindlichkeit sinkt daher auf $\frac{1}{9}$, ab, aber infolge der hohen Halbwertsfrequenz klingt die Empfindlichkeit dieses Bolometers mit wachsender Frequenz so langsam ab, daß sie für hohe Frequenzen diejenige von Bolometern mit geringerem Wärmeverlust übertrifft.

Bei 400 Torr H_2 -Füllung beobachtete man einen Rückgang des Widerstandes um etwa 2,5% und ein Anwachsen des Ausschlags innerhalb von 3 Stunden, der danach zum Stillstand kam. He-Füllung wird daher empfehlenswerter sein.

Will man umgekehrt angenähert die Eigenschaften des Hochvakuumbolometers, d. h. eine hohe statische Empfindlichkeit bei geringem Wärmeverlust erhalten, ohne die Hoch-Vakuum-Schwierigkeiten in Kauf nehmen zu müssen, so empfiehlt sich Xenonfüllung. Zur Orientierung über die Wirkung des Xenons und noch anderer Füllgase bei einem Bolometer von $F = 7 \times 6 \text{ mm}^2$ Fläche und 2,5 mm Reflektorabstand bringt Tabelle 7 eine Übersicht über die statischen Empfindlichkeiten in Prozenten der Empfindlichkeit des Hochvakuumbolometers

Tabelle 7

	Hochvakuum $p = 10^{-3} \text{ Torr}$	X 3 Torr	Ar 3 Torr	Luft 3 Torr	Luft 760 T.	Ne 3 Torr	He 3 Torr	H_2 3 Torr
$\varepsilon(0)$ in %	100	67,5	45	40	36,5	30	20	19,5
V in Volt	0,925	1,37	1,04	2,31	2,53	3,08	4,58	4,66

Tabelle 8

a) Füllgas: Xenon

p (Torr)	0,024	0,29	3,4	44	570
$\nu_{\frac{1}{2}}$ (sec^{-1})	13	19,5	20	13,5	6,5

b) Füllgas: Argon

p (Torr)	0,02	0,3	3,7	48	620
$\nu_{\frac{1}{2}}$ (sec^{-1})	17,5	36,5	43,5	29,5	8

c) Füllgas: Luft

p (Torr)	0,13	4,6	59	760	
$\nu_{\frac{1}{2}}$ (sec^{-1})	39,5	59,5	36,5	9,5	

d) Füllgas: Neon

p (Torr)	0,056	0,27	3,5	45	590
$\nu_{\frac{1}{2}}$ (sec^{-1})	26,5	55,5	109	71,5	16

e) Füllgas: Helium

p (Torr)	0,024	0,30	3,6	46	600
$\nu_{\frac{1}{2}}$ (sec^{-1})	22,5	88	250	120	42

und die Betriebsspannungen, die zur Aufrechterhaltung desselben Widerstandes, d. h. gleicher Übertemperatur, erforderlich sind. Der Widerstand des Bolometers betrug 198Ω . Als Strahlungsquelle diente ein schwarzer Körper von 304°C .

Die vorstehende Tabelle 8 gibt eine Übersicht über die Halbwertsfrequenzen für ein Bolometer von $R = 251 \Omega$ Widerstand, $F = 7 \times 6 \text{ mm}^2$ Fläche, $a = 2,5 \text{ mm}$ Reflektorabstand und $d_M = 0,04 \mu$ Dicke der Membran für eine Reihe von Füllgasen. Für Hochvakuum ist $\nu_1 = 9,5 \text{ sec}^{-1}$.

Die Halbwertsfrequenzen $\nu_1 > 75 \text{ sec}^{-1}$ wurden in der gleichen Weise, wie oben beschrieben, durch Extrapolation aus $\varepsilon(75)$ und $\varepsilon(0)$ berechnet.

§ 11. Optimale Betriebs-Bedingungen und Belastbarkeit

Wenn man ein Bolometer in einen Brückenkreis legt und seine Empfindlichkeit ε definiert durch den Ausschlag des Nullinstrumentes bei irgendeiner Einheitsbestrahlung, so wird man erwarten, daß die so definierte Empfindlichkeit zunächst ansteigt mit der an die Brücke angelegten Spannung. Dies tritt auch in Wirklichkeit ein, doch gibt es für die Spannung einen Maximalwert V_M , bei dessen Überschreitung das Bolometer-Material infolge zu großer Erwärmung Schaden nimmt. Außerdem gibt es aber für die Spannung einen Optimalwert V^* , bei dem die Empfindlichkeit ein Maximum aufweist. Das Auftreten einer solchen optimalen Betriebsspannung ist bekanntlich eine Eigenschaft aller Bolometer. Auch die vorliegenden zeigen sie. Sie erklärt sich im wesentlichen daraus, daß mit wachsender Temperatur der Bolometer-Folie die Strahlungsverluste stark ansteigen und der Temperaturkoeffizient des elektrischen Widerstandes abnimmt.

Die optimale Betriebsspannung der vorliegenden Bolometer ist stark vom Gasdruck im Bolometer abhängig, wofür in Tabelle 9 ein Beispiel gegeben ist.

Tabelle 9

Abhängigkeit der Empfindlichkeit ε (in willkürlichen Einheiten) eines Bolometers von der Spannung V , die an der Folie liegt, und dem Druck p der Luft, in der sich die Folie befindet. Bolometer-Fläche $0,7 \times 0,6 \text{ cm}^2$, Widerstand 190 Ohm , Strahlung mit $2,5 \text{ Hz}$ moduliert

$\begin{array}{c} V \\ p \end{array}$	0,4	0,8	1,2	1,6	1,7	1,8	2,0	2,4 Volt
10^{-3} Torr	60	115	155	170	172	170	161	132

$\begin{array}{c} V \\ p \end{array}$	1,5	2,5	3,0	3,5	3,75	4,0	5,0	6,0 Volt
1,1 Torr	108	158	165	159	150	134	65	—
0,6 Torr	71	107	118	124	126	125	105	69
0,5 Torr	60	93	104	111	113	115	101	70

Ein anderes Bolometer von $0,3 \times 0,2 \text{ cm}^2$ Fläche und etwa 230 Ohm Widerstand ergab (ebenfalls bei $2,5 \text{ Hz}$ Modulations-Frequenz) maximale Empfindlichkeit bei folgenden Werte-Paaren von Druck und Betriebsspannung: $p = 10^{-3} \text{ Torr}$, $V^* = 0,75 \text{ Volt}$. $p = 0,1 \text{ Torr}$, $V^* = 1,75 \text{ Volt}$. $p = 0,75 \text{ Torr}$, $V^* = 2,5 \text{ Volt}$. $p = 2 \text{ Torr}$, $V^* = 2,6 \text{ Volt}$.

Die Frequenz-Abhängigkeit der opt. Betriebsspannung wurde an einem Bolometer von $0,3 \times 0,2 \text{ cm}^2$ Fläche und 127 Ohm Widerstand bei einem Druck von

0,7 Torr Luft untersucht. Es ergab sich beim Übergang der Modulationsfrequenz von 0 auf 50 Hz ein Anstieg der opt. Betriebsspannung von 1,75 auf 2,0 Volt.

Was die maximale Betriebsspannung betrifft, so wurden an den vorliegenden Bolometern folgende Beobachtungen gemacht: Nähert man sich mit der Spannung dem höchstzulässigen Wert, so wird eine ursprünglich glatte Bolometer-Membran faltig.

Das Faltigwerden beginnt — an der Stelle der höchsten Temperatur — in der Mitte der Schicht und breitet sich mit wachsender Belastung über die ganze Fläche aus. Diese Erscheinung läßt sich besonders gut an großen Flächen, z. B. 7×6 und 27×25 mm² beobachten. Die Falten verschwinden zunächst wieder, wenn die Spannung abgeschaltet wird. Erhöht man die Belastung weiter, so verschwinden die Falten nach Abschalten nur noch teilweise: es bleibt eine sichtbare Veränderung zurück. Hier beginnen die Prozesse, die schließlich zur Zerstörung der Schicht führen, nämlich Verschiebungen der Zelluloidgrundlage und Oxydation. Die Schmelzprozesse in der Zelluloidmembran lockern das Gefüge der Schicht. Der Widerstand steigt an und die endgültige Zerstörung erfolgt teils durch Risse in der Membran, teils durch ein Auseinanderfließen der Schicht ohne Riß in der Membran.

Die maximale Betriebsspannung V_M (Volt) und die maximale Belastung $\Phi_M \left(\frac{\text{Watt}}{\text{cm}^2} \right)$, oberhalb welcher die Zerstörung beginnt, zeigt Tabelle 10 nach Versuchen an Bi-Bolometern mit Widerständen von $R = 100$ bis 300Ω . Diese Größen hängen von V_w , d. h. von der Gasfüllung und der Konstruktion ab; der Reflektorabstand ist 2,5 mm.

Tabelle 10

Luftdruck p (Torr)	Bolometerfläche F mm ²	Max. Betriebs- spannung (Volt)	Max. Belastung pro cm ² (Watt/cm ²)
10 ⁻³	3×2	0,11 \sqrt{R}	0,2
0,75		0,245 \sqrt{R}	1
10 ⁻³	7×6	0,185 \sqrt{R}	0,08
0,75		0,44 \sqrt{R}	0,45
0,75	27×25	1,12 \sqrt{R}	0,2

Als gefahrlose Belastung wird man $\frac{1}{2}$, als gefahrlose Betriebsspannung etwa 0,7 der maximalen ansehen dürfen. In der Nähe dieses Wertes pflegt auch die optimale Betriebsspannung zu liegen.

§ 12. Widerstandsrauschen und Empfindlichkeitsgrenze

Das Widerstandsrauschen des Bolometers bestimmt die Grenze seiner Leistungsfähigkeit. Es wurde experimentell auf folgenden zwei Wegen untersucht: 1. wurde das Widerstandsrauschen der Bolometer direkt mit dem von Drahtwiderständen verglichen, welche statt ihm an den Eingang eines rauscharmen Verstärkers gelegt wurden. 2. wurde die kleinste Strahlungsintensität gemessen, die das 3fache der Störspannung hervorrief.

Zu 1. An den Eingang eines Verstärkers wurden möglichst geschützt gegen äußere Störungen nacheinander gelegt:

- a) eine Schicht eines unbelasteten Bolometers,
- b) dieselbe Schicht bei Belastung mit der Betriebsspannung,
- c) ein Drahtwiderstand von 200Ω ,
- d) ein Drahtwiderstand von 1000Ω .

Am Ausgang wurde der verstärkte, mit Trockengleichrichter gleichgerichtete Störstrom des Johnstoneffektes auf einem Photogramm von 3 Minuten Dauer registriert. Man fand, daß das belastete Bolometer ein etwas größeres Widerstandsrauschen zeigte als das unbelastete. Beide lagen nur wenig über dem Rauschen des Drahtwiderstandes von 200Ω und weit unter demjenigen von 1000Ω . In Ausnahmefällen war der äquivalente Widerstand höher oder wuchs mit wachsender Belastung stärker an; solche Schichten wurden als unbrauchbar ausgesondert.

Zu 2. Die kleinste noch nachweisbare Strahlungsintensität wurde auf folgende Weise bestimmt: Die Strahlung eines schwarzen Körpers mit einer Öffnung von $4,9\text{ cm}^2$ und einer Temperatur von 120°C fiel aus 4 m Entfernung auf das Bolometer. Die Strahlung wurde dabei durch einen Drehflügel mit 21 sec^{-1} moduliert. Das Bolometer war über einen Transformator an einen Verstärker angeschlossen. Der verstärkte Wechselstrom wurde durch einen mit dem Drehflügel gekoppelten, mechanischen Gleichrichter umgeformt und einem Drehspul-Galvanometer zugeführt. Die Phasenlage des Gleichrichters relativ zum Drehflügel konnte während des Laufes verstellt werden, bis maximaler Ausschlag erreicht war.

Von der Strahlung des schwarzen Körpers gelangten nur etwa 60% auf das Bolometer, da schätzungsweise mindestens 40% durch Reflexion und Absorption im Verschlussfenster des Bolometers und durch Absorption im CO_2 und H_2O der Luft verloren gingen. Ferner moduliert der Drehflügel die Strahlung nicht genau sinusförmig. Die Amplitude der Grundschwingung betrug nur etwa 80%. Von diesen Verlusten abgesehen, hätte die Strahlungsdichte auf dem Bolometer $9,3\text{ erg sec}^{-1}\text{ cm}^{-2}$ betragen. Man wird sie daher in Wirklichkeit mit etwa $4,6\text{ erg sec}^{-1}\text{ cm}^{-2}$ ansetzen dürfen.

Bei dieser Strahlungsdichte betrug der Ausschlag des Galvanometers das 12-fache des mittleren Störausschlages. Definiert man als kleinste nachweisbare Strahlungsdichte diejenige, die den dreifachen Störausschlag hervorruft, so betrug sie also im vorliegenden Falle $1,1\text{ erg sec}^{-1}\text{ cm}^{-2}$. Da das Bolometer eine Fläche von $0,4\text{ cm}^2$ besaß, betrug die kleinste nachweisbare Strahlungsleistung $0,4\text{ erg sec}^{-1}$.

Ein Vergleich dieses experimentellen Befundes mit den Angaben der entsprechenden theoretischen Formeln für die Leistungsgrenze der Bolometer, wie sie am Ende des § 6 zusammengestellt sind, wurde zur Zeit der Messungen nicht durchgeführt. Ein nachträgliches Einsetzen von annehmbaren Werten in diese Formeln läßt eine etwa viermal größere Nachweisempfindlichkeit des Bolometers erwarten, als sie experimentell gefunden wurde. Dieser Unterschied läßt sich jetzt aus den Protokollen nicht mehr aufklären. Möglicherweise war der Störpegel durch Fremdfelder aus dem Laboratorium erhöht. In dieser wie in mancher anderen Hinsicht wären weitere Messungen erforderlich, die wir hoffen, in absehbarer Zeit wieder aufnehmen zu können.

Frankfurt a. M., Physikalisches Institut der Universität.

(Bei der Redaktion eingegangen am 7. Juni 1950.)

Über die Beschneidung des Rauschspektrums bei Schwingungskreisen

(Eine Ergänzung zu einer Arbeit von R. Feldtkeller)

Von Wolfgang Böer

(Mit 4 Abbildungen)

Inhaltsübersicht

In Ergänzung einer Arbeit von R. Feldtkeller wurde das Rauschen von Schwingungskreisen in Abhängigkeit der Durchlaßbandbreite eines nachfolgenden Verstärkers berechnet. Es wurde gefunden, daß die Größe der Rauschspannung vom Betrage der Dämpfung der Kreise abhängt, und zwar mit wachsender Dämpfung abnimmt. Außerdem wächst die Rauschspannung mit zunehmender relativer Bandbreite. In der nachfolgenden Diskussion ergibt sich, daß es keineswegs immer günstig ist mit möglichst schwach gedämpften Eingangskreisen zu arbeiten. Auch lassen sich für spezielle Probleme Aussagen über das günstigste Verhältnis von Kapazität zu Induktivität machen.

In Ergänzung des von R. Feldtkeller¹⁾ behandelten Problems des Rauschens von Schwingungskreisen erscheint es Verf. von Wichtigkeit, dieses Rauschen bei nachfolgender selektiver Verstärkung zu untersuchen. In jener Arbeit wurde ausgehend von der Rauschgleichung für Widerstände von Nyquist²⁾

$$\overline{u^2} = 4 k T \int R \cdot d\nu \quad (1)$$

für R der Realteil des komplexen Scheinwiderstandes des Schwingungskreises eingesetzt. Eine dann zu erfolgende Integration wurde jedoch über das gesamte ν -Intervall erstreckt.

R. Feldtkeller erhält so für das Schwankungsquadrat einen vom Widerstand und der Induktivität unabhängigen Wert

$$\overline{u_{\text{ges.}}^2} = \frac{k T}{C}. \quad (1a)$$

Dies Vorgehen ist, obgleich es zu einem endlichen Grenzwerte führt, nur dann zu rechtfertigen, wenn nur Frequenzgebiete einer Umgebung der Resonanzfrequenz, die kleiner ist als die Durchlaßbandbreite des nachfolgenden Verstärkers, Rausch-

¹⁾ R. Feldtkeller, Veröff. a. d. Geb. d. Nachr. techn. 6, 2. Folge (1936).

²⁾ H. Nyquist, Physic. Rev. 32, 110 (1928).

anteile liefern; denn in allen praktisch zur Diskussion stehenden Fällen interessieren nur Rauschanteile bei nachfolgender selektiver Verstärkung.

Es wird verständlich, daß diese Bedingung nicht stets erfüllt sein kann, wenn man das Rauschen eines Schwingungskreises als Funktion seiner Dämpfung bei fester Bandbreite des nachfolgenden Verstärkers betrachtet. Nur bei kleiner Dämpfung dürfte als Grenzfall die Formel (1a) gelten, für größere Dämpfungswerte müßte jedoch eine Abhängigkeit von der Verstärkerbandbreite auftreten, da Rauschwerte außerhalb der Durchlaßbandbreite keinen Beitrag zum Rauschen liefern, so daß für das Integrationsintervall eine Verminderung der Rauschwerte resultieren müßte. Es ist hingegen bekannt, daß der aus (1a) berechnete Rauschwert zum Teil sogar größenordnungsmäßig über dem gemessenen Wert liegt, und bereits R. Feldtkeller weist in seiner Arbeit auf diese Tatsache hin.

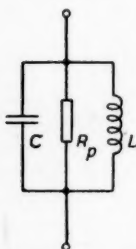


Abb. 1. Schwingungskreis mit Verlusten

Es soll daher in folgenden die Größe des Schwankungsquadrates in einem endlichen Frequenzintervall bei zwei verschiedenen Schwingungskreisen berechnet werden. Da eine exakte Integration zwischen endlichen Grenzen zu unübersichtlichen Resultaten führt, sollen folgende vereinfachende Annahmen gemacht werden:

- a) Das Durchlaßfrequenzintervall möge symmetrisch zur Resonanzfrequenz ν_0 liegen,
- b) Der Betrag der Bandbreite sei klein gegen die Resonanzfrequenz.
- c) Für die Berechnung des Kreises nach Abb. 3 wird die Annahme gemacht, daß $\omega_0^2 R^2 C^2 \ll 1$, oder $R^2 \ll \frac{L}{C}$ ist.

Als Schwingungskreis sei zuerst ein Kreis mit Verlusten untersucht, dessen Ersatzschaltbild Abb. 1 wiedergibt. Es ist dann der komplexe Widerstand

$$\Re = \frac{1}{\frac{1}{R} + j\left(\omega C - \frac{1}{\omega L}\right)}. \quad (2)$$

Der Realteil des Widerstandes ist demnach

$$R_\omega = \frac{\frac{1}{R}}{\frac{1}{R^2} + \left(\omega C - \frac{1}{\omega L}\right)^2}. \quad (3)$$

Unter Einführung von $\frac{1}{R} = \varrho$ und $\frac{1}{\sqrt{LC}} = \omega_0$ folgt

$$R_\omega = \frac{\varrho}{\varrho^2 + \frac{C^2}{\omega^2}(\omega^2 - \omega_0^2)^2}. \quad (4)$$

Aus der Annahme b) folgt die Näherung

$$\omega^2 - \omega_0^2 \approx 2\omega_0(\omega - \omega_0) \quad (5)$$

$$\frac{\omega}{\omega_0} \approx 1. \quad (6)$$

Also ist

$$R_\omega = \frac{\varrho}{\varrho^2 + 4C^2(\omega - \omega_0)^2}. \quad (7)$$

Die Integration erstreckt sich jetzt von ν_1 bis ν_2 mit $\nu_0 - \nu_1 = \nu_2 - \nu_0 = \frac{\Delta\nu}{2}$. Damit wird das Schwankungsquadrat

$$\overline{u^2} = 4 k T \int_{\nu_1}^{\nu_2} \frac{\varrho d\nu}{\varrho^2 + 16 C^2 \pi^2 (\nu - \nu_0)^2} \quad (8)$$

mit $x = \nu - \nu_0$ folgt, da das Integral symmetrisch in x ist

$$\overline{u^2} = 8 k T \int_0^{\frac{\Delta\nu}{2}} \frac{\varrho dx}{\varrho^2 + 16 \pi^2 C^2 x^2} \quad (9)$$

$$= \left[\frac{2 k T}{\pi C} \operatorname{arctg} (4 \pi R C x) \right]_0^{\frac{\Delta\nu}{2}} \quad (10)$$

$$= \frac{2 k T}{\pi C} \operatorname{arctg} (2 \pi R C \Delta\nu). \quad (11)$$

Für $2 \pi R C \Delta\nu \gg 1$ geht dieser Wert in das Resultat von R. Feldtkeller über. Praktisch ist das aber keineswegs erfüllt. Es gilt dann Gl. (11), die für das Schwankungsquadrat wie erwartet eine Abnahme bei Verkleinerung der Band-

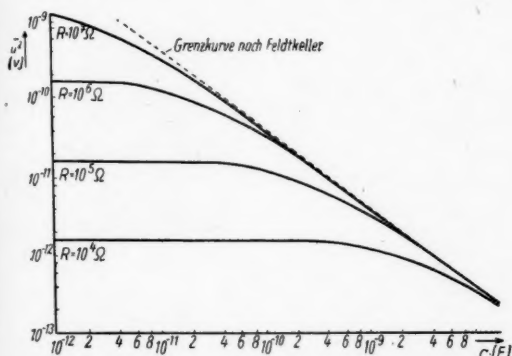


Abb. 2. Schwingkreisrauschen bei Kreisen nach Abb. 1. Bandbreite des nachfolgenden Verstärkers 10^4 Hz

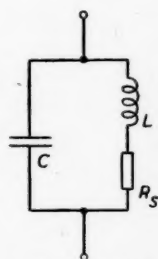


Abb. 3. Schwingungskreis ohne kapazitive Verluste. R_s = ohmscher Widerstand der Induktion

breite bzw. Vergrößerung der Dämpfung zeigt. Aber diese Abnahme erfolgt nicht mehr linear, sondern nach einer Arcustangensfunktion, d. h. für größere Werte des Argumentes wächst die Größe der Schwankung schwächer als bei rein ohmschem Widerstand. Abb. 2.

Für einen Schwingkreis, dessen Ersatzschaltbild durch Abb. 3 dargestellt wird, berechnet sich

$$\Re = \frac{1}{j \omega C \frac{1}{R + j \omega L}}; \quad (12)$$

unter Einführung von $\omega_0 = \frac{1}{\sqrt{LC}}$ ergibt sich für den Realteil

$$R_\omega = \frac{R \omega_0^4}{(\omega_0^2 - \omega^2)^2 + \omega^2 \omega_0^4 R^2 C^2} \quad (13)$$

unter der Berücksichtigung von b) bzw. (5) und (6) folgt

$$R_\omega = \frac{R \omega_0^2}{4 (\omega - \omega_0)^2 + \omega^2 \omega_0^2 R^2 C^2}. \quad (14)$$

Damit wird das Schwankungsquadrat

$$\overline{u^2} = 4 k T \omega_0^2 R \int_{\nu_1}^{\nu_2} \frac{d\nu}{16 \pi^2 (\nu - \nu_0)^2 + 16 \pi^4 \nu^2 \nu_0^2 R^2 C^2} \quad (15)$$

wegen Annahme c) läßt sich dafür schreiben, wenn $x = \nu - \nu_0$

$$\overline{u^2} = 2 \frac{k T \omega_0^2 R}{4 \pi^2} \int_0^{\frac{\Delta \nu}{2}} \frac{dx}{x^2 + 2 \pi^2 \nu_0^2 R^2 C^2 x + \nu_0^2 R^2 C^2 \pi^2}. \quad (16)$$

Da

$$\pi^2 \nu_0^4 R^2 C^2 (\nu_0^2 R^2 C^2 \pi^2 - 1) \approx -\pi^2 \nu_0^4 R^2 C^2 < 0 \quad (17)$$

ist, gilt

$$\overline{u^2} = \left[\frac{k T \omega_0^2 R}{2 \pi^2 R C \pi \nu_0^2} \operatorname{arctg} \left(\frac{2 \nu_0^3 R^2 C^2 \pi^2 + 2 x}{2 \nu_0^2 \pi R C} \right) \right]_0^{\frac{\Delta \nu}{2}} \quad (18)$$

$$\overline{u^2} = \left[\frac{2 k T}{\pi C} \operatorname{arctg} \left(\pi \nu_0 R C + \frac{x}{\pi \nu_0^2 R C} \right) \right]_0^{\frac{\Delta \nu}{2}}, \quad (19)$$

also ist, sofern man berücksichtigt, daß

$$\omega_0 R C \ll \frac{\Delta \nu}{\nu_0} \cdot \frac{1}{\omega_0 R C}$$

ist, was praktisch stets erfüllt ist,

$$\overline{u^2} = \frac{2 k T}{\pi C} \operatorname{arctg} \left(\frac{\Delta \nu}{2 \nu_0^2 \pi C R} \right) \quad (20)$$

oder

$$\overline{u^2} = \frac{2 k T}{\pi C} \operatorname{arctg} \left(\frac{2 \pi L \Delta \nu}{R} \right). \quad (21)$$

Bei konstantem L und $\Delta \nu$ wird das Schwankungsquadrat um so größer, je kleiner R , also je kleiner die Dämpfung ist, so wie es erwartet wurde.

Für $\frac{L 2 \pi \Delta \nu}{R} \gg 1$ geht auch hier Gl. (21) in das Resultat von R. Feldtkeller über; dann liegen praktisch alle anteilliefernden Rauschwerte innerhalb der Durchlaßbandbreite.

Werden in den Gln. (11) und (21) die Argumente der Arcustangensfunktionen so umgeschrieben, daß in ihnen die Dämpfung der Kreise erscheint, die im ersten Falle

$$D_1 = \frac{1}{\omega_0 R_p C} \quad (22)$$

beim zweiten Kreis

$$D_2 = \frac{R_s}{\omega_0 L} \quad (23)$$

beträgt, so werden beide Gleichungen formal identisch, wie man es nach den eingangs geführten Überlegungen erwarten mußte. Es gilt dann

$$\overline{u^2} = \frac{2 k T}{\pi C} \operatorname{arctg} \left(\frac{1}{D} \cdot \frac{\Delta \nu}{\nu_0} \right). \quad (24)$$

Es tritt jetzt allerdings die relative Bandbreite an die Stelle der absoluten, und man wäre versucht, bei höheren Frequenzen geringere Rauschwerte zu erwarten, da $\Delta \nu$ vom nachfolgenden Verstärker bestimmt, nicht mit D in Beziehung steht. Praktisch ist jedoch $\Delta \nu / \nu_0$ mit 10^{-4} größenordnungsmäßig eine Grenze gesetzt, die im gesamten interessierenden Frequenzgebiet etwa konstant ist.

Ob eine Vergrößerung der Dämpfung eine Verbesserung des Verhältnisses Meßspannung zu Rauschspannung bewirkt, auf das es ja in Verstärkeranordnungen ankommt, läßt sich summarisch nicht beantworten. Es wird bei zunehmender Dämpfung eine Verringerung der Rauschwerte resultieren, die bestenfalls der Wurzel aus der Dämpfung proportional ist. Da aber für den Resonanzwiderstand des Kreises

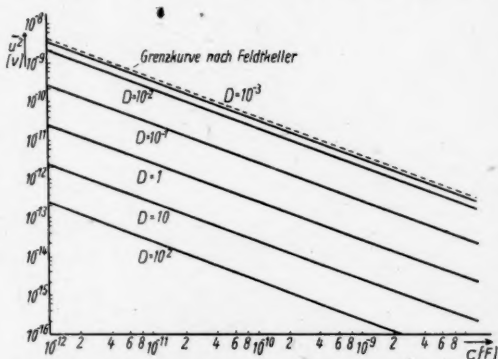


Abb. 4. Schwingkreisrauschen in Abhängigkeit von der Dämpfung des Schwingungskreises bei einer relativen Bandbreite des nachfolgenden Verstärkers von $\Delta \nu / \nu_0 = 10^{-3}$

$$R = \frac{L}{C R_s} = \frac{1}{\omega_0 C} \cdot \frac{1}{D} \quad (25)$$

gilt, kann eine Beeinflussung der Größe der Meßspannung stattfinden, indem diese ebenfalls verringert wird. Dieser Effekt kann dann natürlich die Verbesserung des Meß-, Rauschspannungsverhältnisses überkompensieren. Im allgemeinen wird man daher sagen können, daß in Fällen, in denen eine bestimmte Meßspannung vorgegeben ist, man mit relativ stark gedämpften Eingangskreisen, dort, wo jedoch ein bestimmter Strom vorgegeben ist, besser mit schwach gedämpften Kreisen (um einen möglichst großen Spannungsabfall an ihnen zu erreichen) arbeiten

wird, wobei sich die optimalen Arbeitsbedingungen aus der speziellen Schaltung und den Gln. (24) und (25) errechnen lassen.

Für die Erzielung der notwendigen Dämpfung ist es, wie Abb. 2 zeigt, überdies keineswegs gleichgültig, ob sie durch dielektrische bzw. Induktionsverluste oder z. B. durch hochohmigen Induktionsspulendraht erzeugt wurde.

Von besonderer Wichtigkeit ist das Verhältnis von $L : C$. Zwar wird durch eine Vergrößerung von C eine Verringerung der Rauschwerte erreicht, die aber auch hier bestenfalls nur proportional zur Wurzel aus C erfolgt (gewöhnlich aber bedeutend geringer; vgl. Abb. 2) während diese Vergrößerung von C nach Gl. (25) durch Verringerung des Resonanzwiderstandes des Schwingungskreises unter Umständen eine bedeutend größere Verringerung der am Kreise liegenden Meßspannung bewirkt. Auch hier sind die optimalen Arbeitsbedingungen nur aus dem speziellen Problem zu errechnen.

Berlin, II. Physikalisches Institut der Universität Berlin.

(Bei der Redaktion eingegangen am 14. Juni 1950.)

Verantwortlich

für die Schriftleitung: Prof. Dr. Friedrich Möglich, Berlin-Buch, Lindenberger Weg 74;
für den Anzeigenteil: Ernst Wöllnitz (Arbeitsgemeinschaft medizinischer Verlage G. m. b. H.),
Berlin C 2, Neue Grünstraße 18, Fernruf: 52 20 97. Z. Z. gilt Anzeigenpreisliste Nr. 1
Verlag: Johann Ambrosius Barth, Leipzig C 1, Salomonstr. 18 B, Fernruf: 63 105, 63 781
(1) Paul Dünhaupt, Köthen L 264/54

Lizenz-Nr. 410 / G. N. 446